

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: July 18, 2002, 04:55:36 ; Search time 1761.63 Seconds
(without alignments)
11791.247 Million cell updates/sec

Title: US-09-732-680a-1
Perfect score: 1539
Sequence: 1 attcgacagcagggatcac.....cttacaacaaaaaaaaa 1539

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 13736207 seqs, 6748477542 residues

Total number of hits satisfying chosen parameters: 27472414

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

EST: *
1: em_estba: *
2: em_esthum: *
3: em_estlin: *
4: em_estlun: *
5: em_estlov: *
6: em_estlpi: *
7: em_estro: *
8: em_hlc: *
9: gb_estl: *
10: gb_estl2: *
11: gb_hlc: *
12: gb_gss: *
13: em_gss_hum: *
14: em_gss_inv: *
15: em_gss_pln: *
16: em_gss_vtc: *

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
C 1	400.8	31.9	890	12 CNS013TY	AL103312 Drosophila
C 2	436.8	28.4	448	9 A1542842	SD09326.5
C 3	328	21.3	361	9 AM944588	SD09326.3
C 4	79.6	5.2	891	9 A110117	GH09582.5
C 5	74	4.8	669	9 BB653265	BB653265
C 6	69.4	4.5	641	9 BB626219	BB626219
C 7	67.6	4.4	554	10 BE664308	BE664308
C 8	67.6	4.4	755	9 AM914206	EST345510
C 9	66.8	4.3	479	9 AA636687	vr15f05.r
C 10	66.6	4.3	837	10 B114221	602862588
C 11	65.4	4.2	713	10 BF309251	601890196
C 12	65	4.2	687	10 BB639700	BB639700
C 13	64.8	4.2	654	9 BB626210	BB626210
C 14	64.4	4.2	660	10 BF307986	601894231
C 15	64.4	4.2	806	10 BF305214	601892667
C 16	64.4	4.2	1052	9 AL530299	AL530299
C 17	63.6	4.1	566	9 AA596914	AA596914

18	63.6	4.1	651	9 BB613689	BB613689
19	62.6	4.1	497	9 AA636794	AA636794
20	62.2	4.0	494	9 AA697326	AA697326
21	62.2	4.0	754	12 CNS02R2B	AL120044 Tetradon
22	61	4.0	864	9 A0120692	AU120692
23	60	3.9	619	10 BM492021	BM492021
24	58.8	3.8	495	9 AA692456	vr21e02.r
25	58	3.8	531	10 BJ065441	BJ065441
26	58	3.8	615	10 BJ032622	BJ032622
27	58	3.8	632	10 BJ035634	BJ035634
28	58	3.8	682	10 BG018521	BG018521
C 29	57.8	3.8	856	12 CNS04DRH	AL286118 Tetradon
C 30	57.4	3.7	585	12 A2305495	A2305495
C 31	56.6	3.7	443	12 FR0034247	AL030614 Fugu rubr
C 32	56.2	3.7	564	10 BM035632	BM035632
C 33	56.2	3.5	694	10 B114994	B114994
C 34	54.6	3.5	589	9 BB610036	BB610036
C 35	54	3.5	438	10 BF544638	BF544638
C 36	53.4	3.5	428	9 AV633191	AV633191
C 37	53.4	3.5	773	10 BF261981	BF261981
C 38	53.4	3.5	822	10 BF301720	BF301720
C 39	53.2	3.5	563	10 B1195946	B1195946
C 40	52.8	3.4	976	12 B1149983	ENTOD93TF
C 41	52.6	3.4	622	10 BE382752	BE382752
C 42	52.2	3.4	809	10 BG404575	BG404575
C 43	52	3.4	468	10 BG553005	BG553005
C 44	52	3.4	982	12 CNS03F40	AL241233 Tetradon
C 45	51.8	3.4	556	9 AW721149	AW721149

ALIGNMENTS

RESULT 1
CNS013TY/C
LOCUS
DEFINITION
CNS013TY 890 bp DNA linear GSS 28-JUL-1999
Drosophila melanogaster genome survey sequence sp6 end of BAC
BACN10F18 of DrosBAC library from Drosophila melanogaster (fruit
fly), genomic survey sequence.

ACCESSION
AL103312
VERSION
AL103312.1 GI:5614923
KEYWORDS
GSS.

SOURCE
Drosophila melanogaster

ORGANISM
Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
Pterygota; Neoptera; Endopterygota; Diptera; Brachycera;
Muscomorpha; Ephydroidea; Drosophilidae; Drosophila.

REFERENCE
1 (bases 1 to 890)
Genoscope.
Direct Submission
Submitted (23-JUL-1999) Genoscope - Centre National de Sequencage :
BP 191 91006 Evry Cedex - FRANCE (E-mail : seqref@genoscope.cns.fr
- Web : www.genoscope.cns.fr)

COMMENT
Determination of this BAC-end sequence was carried out as part of a
collaboration with the European Drosophila Genome project (EDGP) -
http://www.edgp.ebi.ac.uk. This Drosophila melanogaster BAC
library (Dros BAC) was made by Alain Billaud at CEPH (Centre
d'Etude du Polymorphisme Humain) with funding provided by a MRC
project grant. The DNA was prepared from embryos by Alain Bucheton
and Genevieve Payan. It has been constructed in the vector
pBelobAC11.

FEATURES
Source
Location/Qualifiers
1..890
/organism="Drosophila melanogaster"
/plasmid="pBelobAC11"
/db_xref="taxon:7227"
/clone_lib="DrosBAC"
/clone="BACN10F18"
/note="end : SP6"

BASE COUNT
205 a 193 c 240 g 198 t 54 others
ORIGIN

Query Match	Similarity	Score	490.8	DB 12	Length 890
Best Local	Similarity 95.3%	Pred. No. 2.6e-87			
Matches 502	Conservative 10	Mismatches 14	Indels 1	Gaps 1	
OY	352	cgatgagagagagagagcgctgtgacacgctacataatgacacatcagcagatc	411		
Db	527	CGATGAGAGAGAGAGAGANNCCTTTGTCAACCGTACATAATAGACATACGAGATC	468		
OY	412	actttgaagtcacagcagagtcgtgaccccccaatcacactcttcaacgycagcgaagt	471		
Db	467	ACTTTGAAGTCCAGCGAGGCTGAGACCCCCAAATCACACTCTTCCACGCGCAGCAAGT	408		
OY	472	ggccctgagtgccggaaccccaagtgatcc-ctcagccaagatgycacatccggtgagtc	530		
Db	407	GGCCGTAGTGCCGGAACCCAGGTGACCTTCCGACCAAGATGGCCACTTCCGGTGATGCC	348		
OY	531	tcacaocgctgtacaagcgcctactgtcgaactacaactgtcacaactgtgcccccaagaca	590		
Db	347	TCCAGCCGCTGTACACAGGCTTACTGGAACTCACTTCTCAACTGTGCCCCACGACAAAGCA	288		
OY	591	gaagctcgaagttgaagatcgcgtcccttgggaccttgaaggctcgtcctcycgagagaagc	650		
Db	287	GAGCTGCAAGTTGAAGATGGGCTCCTGGGGGCTGAAGGTCGTCCTGCCGGAACAGGCAC	228		
OY	651	ggcggagagagagatccctctgacccaagaagactggttaagttacccgagatgggaatcgt	710		
Db	227	GGCGAGAGAGAGTCCCTTGACCCAGAGACCTGGTTTAGTCACCGGAGTGGGAATGT	168		
OY	711	gaagtcgagagacacacttgtcatgcagagactcagcagcgtacatcagagatgaactctac	770		
Db	167	GGACTTCGGACGCCCACTTTDTCAKTCAGSACTACTACGGCTACATGAGATCAATYTTGAC	108		
OY	771	ggctcagcggcgctcctccatgtacaagcgcgctacatacaaccgcgctccctgcatcgt	830		
Db	107	GGCTCAGCGGGCGCTCCTCATATGACAGCGCGCTCATFTACNANCCGGCTCCTCGTTTNT	48		
OY	831	catcctgacccctcagcctcttgctgctgctcccaatgagggcg	877		
Db	47	CTTCYTBASACCTTAGCCTTCTGGCTGCTCCCAATGKTCGGCG	1		
RESULT 2					
LOCUS	A1542842	448 bp	mRNA	linear	EST 23-APR-2001
DEFINITION	SD09326.5prlme SD Drosophila melanogaster Schneider L2 cell culture				
	PORT2 Drosophila melanogaster cDNA clone SD09326 5 similar to				
	G311822: F8a0011822 'ion channel' located on: 2L 21c5-21c5.;				
ACCESSION	A1542842				
VERSION	A1542842.2	GI:13771762			
KEYWORDS	EST.				
SOURCE	Drosophila melanogaster				
ORGANISM	Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;				
	Pterygota; Neoptera; Endopterygota; Diptera; Brachycera;				
	Muscomorpha; Ephydroidea; Drosophilidae; Drosophila.				
	1 (bases 1 to 448)				
REFERENCE	Harvey,D., Broksstein,P., Hong,L., Evans-Holm,M., Su,C., Tsang,G.,				
AUTHORS	Lewis,S. and Rubin,G.M.				
	BDGP/HHMT Drosophila EST Project				
TITLE	Unpublished (2001)				
JOURNAL	On Mar 19, 1999 this sequence version replaced gi:4460215.				
COMMENT	Other ESTs: SD09326.3prlme				
	Contact: Stapleton, M.				
	BDGP				
	Lawrence Berkeley National Lab				
	One Cyclotron Rd, Berkeley, CA 94720, USA				
	Fax: 510 486 6798				
	Email: http://www.fruitfly.org/EST , est@fruitfly.berkeley.edu				
	hit genomic AB003589: arm:2L [301580,604052]				
	estimated-cyto:21B7-21C7: 04/13/2001				
	Plate: SD.93 row: C column: 2				
	High quality sequence stop				

FEATURES	POLYA=NO.	Location/Qualifiers
source	1..448	/organism="Drosophila melanogaster" /db_xref="taxon:7227" /clone="SD09326" /clone_lib="SD Drosophila melanogaster Schneider I2 cell culture pot2" /lab_host="DH5-alpha" /note="Vector: pot2; Site_1: EcoRI; Site_2: XhoI; Sized fractionated cDNAs were directly ligated into POT2. Plasmid cDNA library."
BASE COUNT	115 a 140 c 116 g 77 t	
ORIGIN		
Query Match	28.4%; Score 436.8; DB 9; Length 448;	
Best Local Similarity	98.4%; Pred. No. 1.3e-76;	
Matches 441; Conservative	0; Mismatches 7; Indels 0; Gaps 0;	
QY	7	cacgaggtacatccgaaacaaggcgcgcgtgaacaatgacgacgactcccagaataag 66
DB	1	CAGCAACCGTACATCCGAAACAAAGCGCGCTTACACATGACGACGACTGCCAAGATAAG 60
QY	67	gcaccagttccgcgtctcgtgactgacactactgtctgcaatgctaaatggagatctctt 126
DB	61	GCACCAAGTTCGCGTCTCTGACTGGCAGTACACTGCTGCAATGCTAATGGGATGCTTCT 120
QY	127	atggaggtactactccgtgacgaaggccactgacacgcgcgaccaccaagaagccaatgctc 186
DB	121	ATGGGGCTACTCTCGTGGCAGGCGCAGCTGCACCGCGGACCCCAAGACGCCAATGTC 180
QY	187	aagggccctgatactgcctccacgcgcgcctgttccacgaactacgacagatgtgcagccg 246
DB	181	AAGGGCGCTGATCGCTCCAGCGCGCTGTACAGAACTACGACAGCATGTGACGCGC 240
QY	247	gtgtccaaaggaacccccaagaaagctgtccctgtaaatgtgttcaactcaatagacatc 306
DB	241	GTGTTCCAAGGAACCCCAAGAACTGTCCCTGGGAATGTGTGTCACCTACTATGACATC 300
QY	307	gacgagttgaagcgaagcttgacacccactgtctgtgtaattctcgatggagaacgag 366
DB	301	GACGAGTTGAACGGCAAGCTGACCAACCACTGCTGGCTCGAATCTCGATGGAGAACGAG 360
QY	367	gagcgcggttggaacaacgctcaacatgacaacatcacaagacagatcactttgaagccagc 426
DB	361	GAGCGCGTGTGGCAACCGTCAATATGACAACATCACCGAGATCACTTTGAAGTCCAGC 420
QY	427	gaggtctggaaccccccaatcacactct 454
DB	421	GAGGTCTGGAGCCCCCAATCACACTCT 448
RESULT 3		
LOCUS	AW944588/c	
DEFINITION	SD09326.3prtime SD Drosophila melanogaster Schneider I2 cell culture	361 bp mRNA linear EST 05-JUL-2001
ACCESSION	AW944588	
VERSION	AW944588.1	GI:8122336
KEYWORDS	EST.	
SOURCE	fruit fly.	
ORGANISM	Drosophila melanogaster	
REFERENCE	Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta; Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha; Ephydroidea; Drosophilidae; Drosophila.	
AUTHORS	Harvey,D., Brockslein,P., Hong,L., Evans-Holm,M., Su,C., Tsang,G., Lewis,S. and Rubin,G.M.	
TITLE	BDGP/HMI Drosophila EST Project	
JOURNAL	Unpublished (2001)	

COMMENT Other-ESTs: SD09326.5prime
Contact: Stapleton, M.
BDGP
Lawrence Berkeley National Lab
One Cyclotron Rd, Berkeley, CA 94720, USA
Fax: 510 486 6798
Email: <http://www.fruitfly.org/EST/estefruitfly.berkeley.edu>
Based upon the presence of a XhoI site followed by a run of 14 or more T residues at the beginning of the sequence, this clone was polyadenylated. The resulting Poly-T sequence has been removed. hit genomic AE003589: Drosophila melanogaster genomic scaffold 14200013386046 section 15 of 16, complete sequence.: 03/19/2001
Plate: SD.93 row: C column: 2
High quality sequence stop: 281.

FEATURES
source location/Qualifiers
1. 361
/organism="Drosophila melanogaster"
/db_xref="taxon:7227"
/clone="SD09326"
/clone_lib="SD Drosophila melanogaster Schneider U2 cell culture port2"
/lab_host="DH5-alpha"
/note="Vector: port2; Site_1: EcoRI; Site_2: XhoI; Sized fractionated cDNAs were directly ligated into port2. Plasmid cDNA library."

BASE COUNT 80 a 98 c 102 g 80 t 1 others
ORIGIN

Query Match 21.3%; Score 328; DB 9; Length 361;
Best Local Similarity 94.2%; Pred. No. 4.7e-55;
Matches 340; Conservative 0; Mismatches 21; Indels 0; Gaps 0;

Qy 1151 gcgagtcgacgagcggagagacaaagagatgtagcagcaccgtagagagcgagc 1210
|||||
Db 361 GCGAGTCGACGAGCGAGAAACCAAGAGATGGAGACACCCGACGAGGCGGACG 302
|||||

Qy 1211 agcaagagtcacgtcgcgtggcgatcacaacacccagagtgccggcgcaagggcaac 1270
|||||
Db 301 AGCAGAGATCCAGTCCCTGGGCAATCAACACCCGAGGTGCGGGCCCAAGGCCAACCC 242
|||||

Qy 1271 agtcgacgtggcgctgtgtgcccacgcgctgtagcgcgcatctctctcctcaagc 1330
|||||
Db 241 AGTTCGATTGGGCGATGCTGGCCACCGCGTAGACCGCATTCCTTCCTTCCTCAGCC 182
|||||

Qy 1331 tggccttcctcatctgtgcccacgcgctgtgtaggagatgctcgagatcaagggca 1390
|||||
Db 181 TGGCCTTCCTCATCTCGCCATCAGCTGTCCTGATAGGATGCTCGACACTCAAGGCCA 122
|||||

Qy 1391 catcccaagcagtgcgacactggaactgattgtgcatcttgcatcttcatgtattaatg 1450
|||||
Db 121 CATCCCAAGCCAGTGCACACTTTGACACTAGTTTTCATTTGCGATTTATGATTAAATG 62
|||||

Qy 1451 tgtgtgacactataattattatga tgaacctglatgaataaagacctctg 1510
|||||
Db 61 TGTGTGCAACTTAAATTCTTATATGACCAACCCCTGATAGGATTAAGTCTCTGCC 2
|||||

Qy 1511 g 1511
Db 1 g 1

RESULT 4
LOCUS A110117 891 bp mRNA linear EST 23-APR-2001
DEFINITION GH09582.5prime GH Drosophila melanogaster head port2 Drosophila melanogaster cDNA clone GH09582.5 similar to nacr1pha-96Aa; FBan0006844 'ion channel' located on: 3R 96A2-96A2.; 04/10/2001, mRNA sequence.
ACCESSION A110117
VERSION A110117.2 GI:13757749
KEYWORDS EST.
SOURCE fruit fly.

ORGANISM Drosophila melanogaster
Eukaryota; Metazoa; Tracheata; Hexapoda; Insecta; Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha; Ephydroidea; Drosophilidae; Drosophila.
REFERENCE 1 (bases 1 to 891)
AUTHORS Harvey, D., Brokstein, P., Hong, L., Evans-Holm, M., Su, C., Tsang, G., Lewis, S. and Rubin, G.M.
TITLE BDGP/HMI Drosophila EST Project
JOURNAL Unpublished (2001)
COMMENT On Aug 26, 1998 this sequence version replaced gi:3478441.
Other-ESTs: GH09582.3prime
Contact: Stapleton, M.
BDGP
Lawrence Berkeley National Lab
One Cyclotron Rd, Berkeley, CA 94720, USA
Fax: 510 486 6798
Email: <http://www.fruitfly.org/EST/estefruitfly.berkeley.edu>
hit genomic AE003748: arm:3R [20186991,20443598]
estimated-cyto:96A1-96A20: 04/10/2001
Plate: GH.95 row: G column: 10
High quality sequence stop: 647
POLYA-No.

FEATURES
source location/Qualifiers
1. 891
/organism="Drosophila melanogaster"
/db_xref="taxon:7227"
/clone="GH09582"
/clone_lib="GH Drosophila melanogaster head port2"
/sex="male and female"
/dev_stage="adult"
/lab_host="DH5 - alpha"
/note="Organ: head; Vector: port2; Site_1: EcoRI; Site_2: XhoI; Sized fractionated cDNAs were directly ligated into port2. Plasmid cDNA library."

BASE COUNT 209 a 270 c 218 g 194 t
ORIGIN

Query Match 5.2%; Score 79.6; DB 9; Length 891;
Best Local Similarity 56.7%; Pred. No. 7.2e-06;
Matches 170; Conservative 0; Mismatches 124; Indels 6; Gaps 1;

Qy 325 ctgacacccacgtgtgtatctccatgagagagcgagcgctgtggaaccg 384
|||||
Db 396 CTACACCAACAGCTGTGGCTGAGACAGCAATGCGAGATCAATTAATTAAGTGGATCCC 455
|||||

Qy 385 tcacaattgacaacatacagcagatcaacttgaaagtcacgagagtcgtgaaccgccaa 444
|||||
Db 456 TCGAGATATGGCGCGCTACAGAGCTTATGTCATCCGACATCTGCGCTGCCGAC 515
|||||

Qy 445 atcaaccttcaagcgagagagtgctgtgctgtagccgaaccc-----agtgacc 498
|||||
Db 516 ATCGTCTCTACAAACAAATGCGATGCGAGTACGTGTCACCAACATGACGAAGGCCATC 575
|||||

Qy 499 ctcaagcaagatggtgaccttcggtggtggttccctgacgcgtgtaacgacctactgcaa 558
|||||
Db 576 CTCACATATACCGGCAAGTGTCTGAGACTCCGCCGCATCTTAATCCAGCTGTAG 635
|||||

Qy 559 ctcaacatgctcaactggtcccaagcagcagagctgcaagttgaaatcggtctctg 618
|||||
Db 636 ATTGATGTGCGCTACTTCCCTTCGATCAGCAGACCTGCTTCATGATGATGCTGGCTGTGG 695
|||||

RESULT 5
LOCUS BB653265 669 bp mRNA linear EST 26-OCT-2001
DEFINITION BB653265 RIKEN full-length enriched, adult male hippocampus Mus musculus cDNA clone C630019M18 5', mRNA sequence.
ACCESSION BB653265
VERSION BB653265.1 GI:16487122
KEYWORDS EST.
SOURCE house mouse.
ORGANISM Mus musculus

OY	531	tccagcgcgtgtaacaaggctactggaactcaaatgctcaactcgccccacagaaga	590
Db	451	TCTGTCACATCTTTAAAGACTTACTGGAGATCATGTCCACCACCTTCCCTTGACGACA	510
OY	591	gagctgcaagtgaagatcgctcttg	618
Db	511	GAACTGCAGCATGAAGCTGGGCACTCG	538
RESULT	8		
LOCUS	AW914206	755 bp	mRNA linear EST 25-MAY-2000
DEFINITION	ESR345510 Normalized rat brain, Bento Soares Rattus sp. cDNA clone		
ACCESSION	RCJA881.5'	end, mRNA sequence.	
VERSION	AW914206		
KEYWORDS	AW914206.1 GI:8079880		
SOURCE	EST.		
ORGANISM	Rattus sp.		
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;		
AUTHORS	Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;		
TITLE	Rattus.		
JOURNAL	1 (bases 1 to 755)		
COMMENT	Lee,N.H., Glodde,A., Chandra,I., Mason,T.M., Quackenbush,J., Kerlavage,A.R. and Adams,M.D. Rat Genome Project: Generation of a Rat EST (RST) Catalog and Rat Gene Index Unpublished (1998) Contact: Lee, NH The Institute for Genomic Research 9712, Medical Center Drive, Rockville, MD 20850, USA Tel: (301)-838-3529 Fax: (301)-838-0208 Email: nhlee@tigr.org This clone is available through the ATCC, contact the ATCC, tel#703-365-2700 for further information Seq primer: M13 Reverse.		
FEATURES	location/qualifiers		
source	1..755		
BASE COUNT	192 a 219 c 176 g 168 t		
ORIGIN	/organism="Rattus sp." /db_xref="taxon:10118" /clone="RCJA881" /clone_lib="Normalized rat brain, Bento Soares" /note="Organ: brain; Vector: pUT73pac; Site_1: EcoRI; Site_2: NotI"		
Query Match	4.4%; Score 67.6; DB 9; Length 755;		
Best Local Similarity	53.1%; Pred. NO. 0.0017;		
Matches 170; Conservative	0; Mismatches 144; Indels 6; Gaps		
OY	305	tccagcagttcaacggcgaactgaccaccaactgctgctgcatgctccgatgagagacy	364
Db	84	TGGATGAGTAACCAACCAATGCAATGAAACCAACCTTGCTGTAAGAATAATCTGATGACT	143
OY	365	aggagcgcgtctggcaacgcgtcaaatatgacaacatcaacgacatcaacttgaaatcca	424
Db	144	ACAAGATTGGAATGCAAAACCCCTCTGTGACTCAACAAAGGGGTGAGTTCATGCGTGTCTGCAG	203
OY	425	gcgaggttcggaccccccaatacaactctttaacggcgagagaagtgtgacctgagggcg	484
Db	204	AGAGATCTGGAAACCAACACATCGTACGTACAAACACGCTGATGGGATTTCCAGGTGG	263
OY	485	aa-----accagggtgacccctcagccacagatggcgaacttcggtgagatgccacgccc	538
Db	264	ATGACAAACCAACCAACCTTACTTCAAGTACACAGGAGAACTGACTTGATGCCGCGGCA	323
OY	539	tgtacaacgacctactggaactcaacatcatgctaacttggccccacagaacagcgaagtga	598
Db	324	TCTTTAAGAGCTCAGCAAATTCAGCGTAGCTACTTCCCAATTCGACTTCGAAAACCTGCA	383

[illegible]

Db 69 GAAACAGCAATGGTGCATATACACTTGAATGCAATCCAGATGACTATGAGAGATGAA 128
 Oy 405 gcagataccttgaagctcagcagagctgtgaccccccaatcaaccttcaac----- 459
 Db 129 AAAAATTCATCCCTCTGGAAAAAGATCTGGGGCGGAGCGCTTCTCTATACAAAGCC 188
 Oy 460 -ggcagcagaagctgagcagctgagcagcagcagcagcagcagcagcagcagcagcagc 518
 Db 189 AGAGCGGACATTTGGCATTTGCAATTCACCAAGSTGCTCGTGGAGTACACCGGACACAT 248
 Oy 519 ccagctgagctcctcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 578
 Db 249 CACCTGACAGCCGACCATCTTAAAGCTATGTGAGATCATTTGACATCACTTCC 308
 Oy 579 ccagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 638
 Db 309 CATCATGATGAGCAAGACTCAGCATGAGCTGGGACCTGAGACTATGACGGCTGTGTGT 368
 Oy 639 ggaagacagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 698
 Db 369 GGACATTTAAACCGGAAATGACACAGGC---GACCTGATGACTTCACTGAGAGCGGGGA 425
 Oy 699 gtgggaacatcgctgagcagcagc 723
 Db 426 GTGGGTGATCAAGAACTCGGGGC 450

RESULT 10
 B114221 837 bp mRNA linear EST 26-JUN-2001
 LOCUS 602862588P1 NIH_MGC_17 Homo sapiens cDNA clone IMAGE:5021829 5',
 DEFINITION mRNA sequence.
 ACCESSION B114221
 VERSION B114221.1 GI:14565122
 KEYWORDS EST.
 SOURCE human.
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 REFERENCE 1 (bases 1 to 837)
 AUTHORS NIH-MGC http://mgc.nci.nih.gov/.
 TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
 JOURNAL Unpublished (1999)
 COMMENT Contact: Robert Strausberg, Ph.D.
 Email: cgabs-remail.nih.gov
 Tissue Procurement: ATCC
 CDNA Library Preparation: Ling Hong/Rubin Laboratory
 CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LMNL)
 DNA Sequencing by: Incyte Genomics, Inc.
 Clone distribution: MGC clone distribution information can be
 found through the I.M.A.G.E. Consortium/LMNL at:
 http://image.llnl.gov
 Plate: LCM1840 row: p column: 22
 High quality sequence stop: 696.
 Location/Qualifiers
 1..837
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /clone="IMAGE:5021829"
 /clone_1id="NIH_MGC_17"
 /tissue_type="rhabdomyosarcoma"
 /lab_host="DH10B (phage-resistant)"
 /note="Organ: muscle; Vector: pORF7; Site_1: EcoRI;
 Site_2: XhoI; cDNA made by oligo-dT priming.
 Directionally cloned into EcoRI/XhoI sites using the
 following 5' adaptor: GGCACGAG(G). Size-selected >500bp
 for average insert size 1.8kb. Library constructed by
 Ling Hong in the laboratory of Gerald M. Rubin (University
 of California, Berkeley) using ZAP-cDNA synthesis kit
 (Stratagene) and Superscript II RT (Life Technologies)."
 BASE COUNT 204 a 232 c 213 g 188 t
 ORIGIN

FEATURES

source

BASE COUNT

ORIGIN

Query Match 4.3%; Score 66.6; DB 10; Length 837;
 Best Local Similarity 50.7%; Pred. No. 0.0027;
 Matches 190; Conservative 0; Mismatches 179; Indels 6; Gaps 1;
 Oy 297 catagacatcagcagctgtaacgcagcagcagcagcagcagcagcagcagcagcagcagc 356
 Db 240 CATCAATGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 299
 Oy 357 gagagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 416
 Db 300 GGTGATTTACAACTTAAATGGAATCCAGATGATGATGATGATGATGATGATGATGAT 359
 Oy 417 gaagtcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 474
 Db 360 TCCTTCAACAGATCTGGCCGCCGACCTTGTCTTATTAATGATGATGATGATGATGAT 419
 Oy 475 ----ctgagtcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 530
 Db 420 TGCTATTTGTCATGATTCACCAAGTCTCTCGAGTACATGCGCCATCAGTGGACAC 479
 Oy 531 tccagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 590
 Db 480 TCCACCATCTTTAAAGCTATGATGATGATGATGATGATGATGATGATGATGATGAT 539
 Oy 591 gagctgcaagctgtaagctcgtcctggtggtggtggtggtggtggtggtggtggtggt 650
 Db 540 GAACGTGACATGAGAGCTGGGACCTGAGCTTACGACGCGCTGTGTGTGTCATTCACCC 599
 Oy 651 ggcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 665
 Db 600 GGAAGCGACACAGCC 614

RESULT 11
 BF309251 713 bp mRNA linear EST 21-NOV-2000
 LOCUS 601890196P1 NIH_MGC_17 Homo sapiens cDNA clone IMAGE:4124058 5',
 DEFINITION mRNA sequence.
 ACCESSION BF309251
 VERSION BF309251.1 GI:11256573
 KEYWORDS EST.
 SOURCE human.
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 REFERENCE 1 (bases 1 to 713)
 AUTHORS NIH-MGC http://mgc.nci.nih.gov/.
 TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
 JOURNAL Unpublished (1999)
 COMMENT Contact: Robert Strausberg, Ph.D.
 Email: cgabs-remail.nih.gov
 Tissue Procurement: ATCC
 CDNA Library Preparation: Ling Hong/Rubin Laboratory
 CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LMNL)
 DNA Sequencing by: Incyte Genomics, Inc.
 Clone distribution: MGC clone distribution information can be
 found through the I.M.A.G.E. Consortium/LMNL at: image.llnl.gov
 Plate: LCM1009 row: p column: 23
 High quality sequence stop: 711.
 Location/Qualifiers
 1..713
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /clone="IMAGE:4124058"
 /clone_1id="NIH_MGC_17"
 /tissue_type="rhabdomyosarcoma"
 /lab_host="DH10B (phage-resistant)"
 /note="Organ: muscle; Vector: pORF7; Site_1: EcoRI;
 Site_2: XhoI; cDNA made by oligo-dT priming.
 Directionally cloned into EcoRI/XhoI sites using the
 following 5' adaptor: GGCACGAG(G). Size-selected >500bp
 for average insert size 1.8kb. Library constructed by

FEATURES

source

BASE COUNT 204 a 232 c 213 g 188 t
 ORIGIN


```

Location/Qualifiers
1..654
/organism="Mus musculus"
/strain="C57BL/6J"
/db_xref="taxon:10090"
/clone="933016516"
/clone_1kb_RIKEN_full-length enriched, adult male
diacephalon
/sex="male"
/tissue_type="diacephalon"
/dev_stage="adult"

```

CDNA Library Prepared by: Ling Hong/Rudin Laboratory
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)

Fri Jul 19 08:26:18 2002

us-09-732-680a-1.rst

Page 11


```

/codon_start=1
/protein_id="CAC37851.1"
/db_xref="GI:13939668"
/translation="MTTPKIKAPVSGPGVLLLOMLMGLMLGLTSVPGATATADRK
NANKVALDRILHAGLFETVDSVDVOPFGPTPNVSELVNVIYIDDELNGKITTCMLN
LRMDEERVNQSOYDINOTITLKSEVNPQDITLFGNDEGCLMAEPQVLTSHDGRK
IMPRAVYATVAYCELMMLNMPHDKQSCKLTGSGWLKVVLPENGTARGESLDDHDLVQSP
EMELVDSRAHFVSODYDYGMEYTLTAQRSSMYAVITVPASCVITALSAFMLPMPH
GEGEIMTGNLLIIVIAFIMFAQLLEPVLINPEPLVLTSTLYLSVSIYIVLVIV
YLGKIKHKKRLPEALRKTLGLHGFAGTLLWLLFESTVTGSOAKETEMHEPVEADEDE
SSPLGINHTVEYPGKANKQDFMALLATATVDRISFVSFLAFITLIRCSV"

```

Query Match	100.0%;	Score 1539;	DB 6;	Length 1539;
Best Local Similarity	100.0%;	Pred. No. 2.7e-256;		
Matches 1539;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;

1	atcggcaaggaggtacatccgaacaagaagcgctcgaaacaattgcagactcccaag	60
Db	1 ATTGGGCGACGAGGGTACATCCGAATCAAGGCGGGCTGAACAAATGAGACGACTCCCAAG	60
Qy	61 ataaagaccacagcttcccggtccctggaactgcacactatgtctgcaaatgtcaatgaggatg	120
Db	61 ATAAAGGACACAGTTTCCGGTCTCGGACTGCGCACTACTGCTGCAAAATGCTAATGGGGATG	120
Qy	121 ctctctatggyggctgacttccgtgcccagggccactgcccacgcgycgagccccaagaagcc	180
Db	121 CTTCCTTATGGGGGTGACTTCCGTGCCAGGCGCCACTGCCACCGGGAGCCCAAGAACGCC	180
Qy	181 aatgtcaagggccctgcatcgctctccaaagccgcgctgtgtcaagaaactgacagcagatgtg	240
Db	181 AATGTCAAGGGCCCTGGATTCGGCTTCCACGCGGGCTGTGTCAAGAACTACGACACGATGTG	240
Qy	241 cagccggtgtttccaaaggaacccccacgaacggtgtcccttgtaaatgtgtgcactctacata	300
Db	241 CAGCCGGTGTTCCAAAGGAACCCCCACGACGCTGTCCCTGGAAATGTGTGCTACCTTACATA	300
Qy	301 gacatcgaacgagattgaaacggaagctgaccacccactgtcgtctgtaattcttcgaatggaga	360
Db	301 GACATCGACGAGTTGGAACGGCAAGCTGACCAACCCACTGTGTGGTGAATCTCCGATGGAGA	360
Qy	361 gaacgaagagcgcggtgtggtgcaacggtccacaatagacaatacaacgcagatcaacttgaag	420
Db	361 GACGAGGCGCCGCTGTGGCAACCTGTCAAAATATGACAACTATCAACGCGATCACTTTGAA	420
Qy	421 tccagcgagaggtctgaccccccaatacacactcttcaagcggaacgaaggttgacctgagt	480
Db	421 TCACGCGAGGTCGTGAGCCCCCCCAATATACACTCTTCAACGCGCGNCAAGGTGGCCTGATG	480
Qy	481 gcccgaaacccaggtgacccctcaagcgaacgatgagccaacttcsgtltgatlgtcctcaagccgtg	540
Db	481 GCCGAACCCAGGTGACCCCTCAGCCACGAGTGGCCACTTCGGGTGAGTGCCTCCAGCGGTG	540
Qy	541 taacagcgctactctgcgaactacaatgctcaactgtgcccaacgaacgaagcgtgcaag	600
Db	541 TACACGGGCTACTCTCGAACTCAACATCTCTCAACTGGGCCCAACGACCAAGCAAGACTGTCAAG	600
Qy	601 ttgaagatccggtctctcttggggggtcgttaaaagtctcttgcgggaagaaaggcagaggaaga	660
Db	601 TTGAAGATCCGGCTCTCGGGGCTTAABAGTCTCTTGCGGGAAGAAAGGCAAGGAGAGGA	660
Qy	661 gagtcaccttgacacagacagactgtgttcagtcacacggagtgaggaaatcgtgtgactccgga	720
Db	661 GAGTCCCTTGACACAGACGAGACTGTGTTCAGTCAACGGGAGTGGGAAATCGTGTGACTCCGA	720
Qy	721 gccacattgtcagtcagagactactacagctacatggaagtaacactgtacggtctcagcgg	780
Db	721 GCCACATTGTTCAGTCAAGAGACTACTACGGCTACATGGAATGAACTGTGACGGGCTCAACGG	780
Qy	781 cgctctccatgatacagggccggtcatctatacaacccggtctcgtgcatcgtgtcatctcggcc	840
Db	781 CGCTCTCCATGATACAGGGCCGGTCTATCTATACAACCCGGTCTCGTGCACTCGTGTGCTCGG	840

Db	781	CGCTCTCCTCAGTACACGAGCGCGTCAATCTACACACCGCGGCTCTGCATGATCTCTCGGAC	840
OY	841	ctctcagcccttcctggtcgtcctccccaatggcggcgagaaatcatatcatcaagcgctg	900
Db	841	CTCTCAGCCCTTCTGGCTCTGCCTCCACATATGGCGCGAGAAATCATATCATCAACGGCTGG	900
OY	901	ctcaccatcgtgatcgcgcgccttctctaatgtatcttcgcaccagctctcctcgaatgctgtcc	960
Db	901	CTCATCATCTGATATGGCGGCTTCTCTCATGTACTTTCGCCCACTCTCTGCAATGCTGTACC	960
OY	961	aacaaatactccactgtgtgtaaatctctctacagcaccaagcctgtgtatcttgacgtctcc	1020
Db	961	AACAAATACCTCCACTGTGTGTAATCTTTCTACACACACAGCAGCTGCTGTATCTGAGCTGTACC	1020
OY	1021	accatcgtctcgaaggtctctatgtcttctacctgtgcacagcgaagcacaaagagcgctccg	1080
Db	1021	ACCATCTGTCTGAGAGTCTGTAGTTCTTACTCTGGCCACAGGCAACAGCAGGGCGCTTGCCG	1080
OY	1081	gagcgctcgagaaagctctctgcacggagacacctgggcacgltgctctcctcagttctc	1140
Db	1081	GAGCGCTCGAGAAAGCTCTCTGCACGGGACCTGGGACAGCTGCTCTCTCTCGGTTC	1140
OY	1141	agcacacctgtgcgagtcgcagagcggagaaagacaaagatatgagcagacaccgtacag	1200
Db	1141	AGCACACTGTGCGAGTCTCCAGCGCGAGAAAGAACCAAGAGATGAGAGACACCCGTACGAG	1200
OY	1201	gagcgcgagcagcagagagttccagtcgcgtgtggcatcaaacacacccgaagltgcggcgcc	1260
Db	1201	GAGCGCGAGCAGAGCAGAGAGTCCGTCCGCTGGGCACTAACACACACCGAGAGTGGCGGGCGCC	1260
OY	1261	aagcgcaaccagttctgacttggcgctgtctgtggaccacgcgcgtggaccgaattctcctgtt	1320
Db	1261	AAGCGCAACCAATTTGACTGGGCGCTGCTGGCACCGCGCTGGACCGCATTTCTTCGTT	1320
OY	1321	tcctcacaacctgtgcctctcctcaatctcttgccacatcagtgctcgcgtgtagagatgcgcga	1380
Db	1321	TCCCTTCAAGCTTGCGCTTCTCTCTCATTTCTGGCCATCAGAGTGTCCGTGTAGGAGATGCTCAGAG	1380
OY	1381	ctcagaagccacatcccaaacccagtcgcgaactctgaactagtttgcaatttgcattcat	1440
Db	1381	CTCAAGGCGCACATCCCAAGCCAGTGCAGCATCTGTGAATAGTTTTCATTTGGCATTTTCAT	1440
OY	1441	gtatttaattgtgtgtgcgaacttaattatttaattgaatgagacctgtatagaataaag	1500
Db	1441	GTAATTAATGTGTGTGCGCAACTTAATAATTAATTAATGATGAGACCTCGATGGAATTAAG	1500
OY	1501	gacctcgcgcgaatgtctgtctacaaaaaanaaaaaaa 1539	
Db	1501	GACCTCTGCGCAATGTCTCTCTTACAAAAAANAANAANA 1539	
RESULT 2			
AY005148			
LOCUS	AY005148	1421 bp	linear
DEFINITION	Drosophila melanogaster	nicotinic acetylcholine receptor beta	INV 01-AUG-2001
ACCESSION	AY005148	subunit mRNA, complete cds.	
VERSION	AY005148.1	GI:15072330	
KEYWORDS			
SOURCE			
ORGANISM			
	fruit fly,		
	Drosophila melanogaster		
	Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;		
	Pterygota; Neoptera; Endopterygota; Diptera; Brachycera;		
	Muscomorpha; Ephydroidea; Drosophilidae; Drosophila.		
REFERENCE			
AUTHORS	Grauso, M. and Sattelle, D.B.		
TITLE	Drosophila melanogaster nicotinic acetylcholine receptor betas		
	subunit (nAcRbeta-21C gene)		
JOURNAL	Unpublished		
REFERENCE	2 (bases 1 to 1421)		
AUTHORS	Grauso, M. and Sattelle, D.B.		
TITLE	Direct Submission		
JOURNAL	Submitted (18-JUL-2000) MRC-FSU Human Anatomy and Genetics,		

TITLE Direct Submission
JOURNAL Submitted (30-Oct-2001) Berkeley Drosophila Genome Project,
Lawrence Berkeley National Laboratory, One Cyclotron Road,
Berkeley, CA 94720, USA

COMMENT Sequence submitted by:
Berkeley Drosophila Genome Project
Lawrence Berkeley National Laboratory
Berkeley, CA 94720

This clone was sequenced as part of a high-throughput process to sequence clones from Drosophila Gene Collection 1 (Rubin et al., Science 2000). The sequence has been subjected to integrity checks for sequence accuracy, presence of a polyA tail and contiguity within 100 kb in the genome. Thus we believe the sequence to reflect accurately this particular cDNA clone. However, there are artifacts associated with the generation of cDNA clones that may have not been detected in our initial analyses such as internal priming, priming from contaminating genomic DNA, retained introns due to reverse transcription of unspliced precursor RNAs, and reverse transcriptase errors that result in single base changes. For further information about this sequence, including its location and relationship to other sequences, please visit our Web site (<http://fruitfly.berkeley.edu>) or send email to cdna@fruitfly.berkeley.edu.

FEATURES
Location/Qualifiers
1..1646
/organism="Drosophila melanogaster"
/db_xref="taxon:7227"
/map="21C5-21C5"
1..1646
/gene="nacrbeta-21C"
/note="alignment with genomic scaffold AE003589"
/db_xref="FLYBASE:FBgn0031261"
37..1020
/gene="nacrbeta-21C"
/note="Longest ORF"
/codon_start=1
/db_xref="FLYBASE:FBgn0031261"
/product="SD09325p"
/protein_id="AAL28170.1"
/db_xref="GI:16769902"
/translation="MTTPKIKAPVSGPLLELLQMLMGLMTSPVATADPK
NANVGLRLHAGLENTNDSDVQPFQCTPNVSLGVVYIDIDELGKLTTHWLN
LRWDEREWMQPSQDNTDTITLKSSEVTPQITFENDEGLAETQVTSIHDHFR
WMPAYVAYCELMNLMHPDKOSCKIKTSGMGLKVVLPENGTAGEISLDDDLVOSP
EMELVDSRAHVSQDYRYGMYETTLTAORSSMTYAVITPASCYITLILSWFLPPHM
GGEKTMINGLITVIAFLMTFAQLPLVLSNMTPLVGCSPLEPHQSSSN"

BASE COUNT 390 a 506 c 428 g 322 t
ORIGIN

Query Match 87.4%; Score 1344.4; DB 3; Length 1646;
Best Local Similarity 90.8%; Pred. No. 1.1e-222;
Matches 194; Conservative 0; Mismatches 31; Indels 121; Gaps 1;

QY 7 cagcagggatccatccgaacaaaggcgctgaacatgaagcagctcccaagataag 66
DB 1 CAGCACCGTACATCCGAACCAAGCGCGCTGACACATATACACAGCTCCCAAGATAAG 60
QY 67 gacacagttccgctctcgtactgactgcaactactgctgcaaatgcttaattggagatgtcttct 126
DB 61 GCACCAAGTTCCGCTTCCTGACTGCACATCTGCTGCAAAATGCTATAGGGATGCTTCTT 120
QY 127 atggagcgtgacttcgtgcaagcgcaactgcaacgcgagaccccaagaagcccaatgtc 186
DB 121 ATGGGAGCTGACTTCGTGCGACGCGCACATGCGACCGGACCCCAAGACGCCAATGTTC 180
QY 187 aagcgcctgtagctcgtccacgacgcgctgttccacgaactagacagcagatgtgacgcg 246
DB 181 AAGGCGCTGAGATCGCTCCACGCGCGCTGTACAGSAATATAGACACGATGTGACGCG 240
QY 247 gtgttccaaagaaaccccaagaaagtgctcctgtgaaatgtgtgtcaactacatagacatc 306
DB 241 GTGTCCAAAGAAACCCCAAGAAAGTGTCTCCGTGGAATGATGTGATCACCATACATGACATC 300

QY 307 gacgagttgaacggaagctgacccaccctgtgtgtatctccgatgtgagagacgag 366
DB 301 GACGAGTTGACGGCAAGCTGACACACCCACTGCTGGCTGAATCTCCGATGGAGAGCGAG 360
QY 367 gaggcggtgtgcaaacgctcaacaatatgacaacatcaacgagatcaactttgaatccagc 426
DB 361 GAGCGGCTGTGGCAACCGTCACAAATATGACAACATCACACCAATATCATTTGAATGCCAGC 420
QY 427 gaggcttgaccccccaaatcacactctctcaacgagcagagagtggtccctgtgacgaa 486
DB 421 GAGGCTTGACACCCCAATTCACCTCTTCAACGGCGAGAGAGGGGCTGATGGCCGAA 480
QY 487 acccagtgtagcctcaagcagatgagcagcttcggtgtatgctccacgacggtgtacag 546
DB 481 ACCCAGTGACCCCTGACGACGATGAGCGCCACTTCCGATGATGCTCCACAGCCGTGACAG 540
QY 547 gctactgtgaaactcaactgtctcaactgtgcccacgaaagcagaagctgtgaagtgaag 606
DB 541 GCTACTGTGGAACCTCAACATGCTCAACTGTGCCCCACGACAAACAGCTGCAAGCTGAAG 600
QY 607 atcgcgtccctgggacctgaagctcgtccctgcggaagacgagcagcagagagatcc 666
DB 601 ATCGGCTCTGGGGCTGTGAAGTCTGCTCCGCGAGAAAGCGACCGCAGAGAGATGCC 660
QY 667 ctgacacgacgacacctgtgtcagtcacccgagctgtggaatcgtgtgactcgcgagccac 726
DB 661 CTTGACCAACGACGACCTGTGTTACACGCGAGTGGAAATGCTGAGACTCGCAGACCCAC 720
QY 727 ttgtgcaatcagagactactagcgtctacatgtgagatcaactgtgagcgtcgaagcgctcc 786
DB 721 TTGTGACGACGAGACTACTACGGCTACATGAGTACACTGTGAGGCCGCCAGCGGCTGCC 780
QY 787 tccatgtacacgagcgctcatctacacacccgctcctcgtcatcgtcatcctgtgacctca 846
DB 781 TCGAGTACACGCGCGGTATCTACACACCGCGCTGATGATGATGATGATGATGATGATGATG 840
QY 847 gctctctgctgctctcccaatgagcgagcagagatcatatgataaaggcctgtcatc 906
DB 841 GCTCTCTGCTGCTCTCCACCATGAGCGCGCGGAGAGATCATATATCAAGCGCTGCTCATC 900
QY 907 atcgtgacgcgcgtctctcctatgtactgtgcccagctcctcgtcagctgtgtccaaacat 966
DB 901 ATCGATGATGCGCGCTCTCTCTCATGTACTGTGCGCCAGCTCTCCAGTGTGCCAACAT 960
QY 967 actccactgtgt----- 978
DB 961 ACTCCACTGTGTGGTGAAGTCTCCCATTTGCATTTCATCAGTTCATGTAATAAGTAG 1020
QY 979 ----- 978
DB 1021 ATGCTAAGTACCCCTAAACATCGCAGAAATATTCGCACTTACCGACACCTCACCGGGCAT 1080
QY 979 -----gtaattctacagcagcagctgtcgtatctgagcgtctccacat 1025
DB 1081 TCCCTAATATGAGTAATCTTCTACAGCACACGCTGCTGTATCGAGCGTCTCACCAT 1140
QY 1026 cgtcgaagttcagttctgtactgtgccaaggaagcaacaaagagcgctgtcggaagc 1085
DB 1141 CGTGAGGTTCTAAGTCTGTACTGTGCGCACAGGCAAGCAACAAGAGCGCCCTGCGGAGGC 1200
QY 1086 gctgagaaagctgtcgtcagcgacctgtgagcagcgtgtgtcgtcgtcgtgtgtgagac 1145
DB 1201 GCTGAGAAAGCTGTGTGACGCGGACCTGTGGCACGTGGCTGCTGTGTGGGTGTTAGCAC 1260
QY 1146 cactgcaagtcgcagagcgagaaagaaagatggaagcagacccgtacagaaagc 1205
DB 1261 CACCGCGAGTGTGCGAGGCGGAGAAACCAAGAGATGTGACGACCCGTACGAGAGAGC 1320
QY 1206 ggaagcagcagagatccagtcgtgtggtacatcaacacacagagatgtgccggtgccaagc 1265
DB 1321 GACGAGCAGAGATCCAGTCCGCTGGGATCATCAACACACGAGGTGCCGGGCGCCAAAGC 1380

Query Match	Best Local Similarity	Match	Score	DB	Length
Query Match	86.4%	Best Local Similarity	99.8%	Pred. No. 4.7e-220;	DB: 3; Length 1334;
Matches 1331;	Conservative	0;	Mismatches 3;	Indels 0;	Gaps 0;

QY	39	aacaaatgaacgaacgaactccccaagataaagacacgaftlltcggltcctgtaactgtaactact	98
Db	1	AACATATGACGACGACTCCCAAGATAAAGGACACAGTTTCCGGTCCGTGACATGCCACTACT	60
QY	99	gctcaaatgtctaaatggggaatgctctcttaatggggcgtacttcctcgccgaagggccaatgac	158
Db	61	GCTGCAATGTCTAATGGGGATGCTTCTTATGGGGCTGACTTCCGTGCCAGGCGCACATGC	120
QY	159	caccgcygaccccccaagaacgcacaaatgltcaaggcccttgaatctgcctctcaacgcgcgctgt	218
Db	121	CACCGCGGACCCCAAGAAGCCCAATGTCAAAGGCCCTTGATGCGCTTCACAGCGCGGCTTT	180
QY	219	caacgaactaagaacgaatgttgcagccggtgttccaaggaaaccccccaagaacgtgtccct	278
Db	181	CACGAACACTACACACGCGCATGTGACGGCGGCTTCCAAAGGAAACCCCCAGAACGTGTCCCT	240
QY	279	gaaatgtgtgtcaacacaaatagaacttgaacgtttgaacacgaagcttgaacacacacatg	338
Db	241	GGAAATGGTGTCTACCTACTATTAACATCGACGAGTTGAACGGCAAGCTGACACCCACTG	300
QY	339	ctgctgtcaatcccgacatgagagagacgaagagcgcgtgtgtgcaacccgttcaacataatga	398
Db	301	CTGGCTGAATTCGGATGGAGAGACGAGGAGCGCGTGTGGCAACCGTACAAATATGACAA	360
QY	399	catcaacgaatcaacttgaatgtccagcgaaggtctgagaccccccaatcaactctcaaa	458
Db	361	CATCACGCGAGATCACTTTGAAAGTCCAGCGAGGCTGTGGACCCCCCAAACTACACTTTCAA	420
QY	459	cggcgacgaaggtgagccttgaatggccgaaaacccagagacccctcaacgaacgtltgacatt	518
Db	421	CGGCGACGAAGTGGCTGTATGGCCGCAAAACCCAGGTGACCCCTTCACCCACAGATGGCACTT	480
QY	519	ccggtgtgactcctccacgcgtgtgacaacgcgtactctcgaactcaacatgtctcaacttggcc	578
Db	481	CCGGTGATGTGCTTCACACCGGTATCAAGGCGTACTGCAATCAACATGTCTCAACTGGCC	540
QY	579	ccaacgaacgaagacgtctgaagtltgaagatcggctctctgtgggccttgaagttcgtctgc	638
Db	541	CCACGACAAAGCAGACTGTGCAAAATTGAAAGATCGGCTCTTGGGGCTGAAAGTGTCCTTGCC	600
QY	639	ggagaaacggacccggcaggaagaagaatcccttgaacacacgaacccgtgttaagttaacccga	698
Db	601	GGAGAACGCGACGGCGAGAGAGAGATCCCTTGACACGACGACGTGCTTCACTCAACCGGA	660
QY	699	gttgggaatatctgtagactcgcgagcccaacttltgcaatcaagagactactaacygtataatga	758
Db	661	GTGGGAATTCGTGACACTCGGAGGCCACCTTGTTCATCACTAGACACTACTACGCGTACATGGA	720
QY	759	gtacaactctgaacgcctcagcgcgtctcctcaatgtataacgcgcgtatcatacaacccgc	818
Db	721	GTACACTTGTGACGGCTCAGCGGCGCTCTTCATAGTACAGCGCGTTCATCTACACACCGC	780
QY	819	gtccctgaatgttaactctggccctctcaagcttctgtcgtcctcccaaatggggggga	878
Db	781	GTCTGCAATCTTCATCTTGCGCCCTCTTAGCTTCTTGCGCTCTCCCAATAGGGGGGA	840
QY	879	gaagatcatgtatcaacgcgcctgtcatcatcgtatcgcgcgcgcctctccatgtactctgc	938
Db	841	GAAGATCATGTATCAACGGCTGCTCATCACTCGATGCGCGCTTCTCTATATTAATTGGC	900
QY	939	ccaagctcctgcagtgctgtgtcaacaatatcctcaacttgtgttaattcttacaagcaccag	998
Db	901	CCAGCTCTCTGCAATGCTGTGTCAAACAATACTCCACTTGGTGAATCTTCTACAGCACAG	960
QY	999	ccctgctgatacttgaggtgtcccaacatcgttcgaagttccagtctgttaacttggccaagg	1058
Db	961	CTGTGTTATTTGAGCGCTTTCACACATCGTCGAGGTTTCAATTCTGTATCTGGCCACAGG	1020
QY	1059	caagacaagaagggcgtctgcggaagcgctgtgaanaagctgtctgcaacggcgaaccttggagac	1118
Db	1021	CAAGACAAGAAGGGCGCTGCGCGAGGCGCTGAGAAAGCTGTTGTGCACAGGGCACTGTGGGAC	1080
QY	1119	gttgagctgtctgtcgtgtgttcaagcaacacttggcgaatgtcgaacgagcgaagaacccaaga	1178

	Db	1081	GTGGGCTGCCTCCTCGTTCATTACAGCACCCGCGCATGTGCAGGGCGAGAACAACCAAGGA	1140
	OY	1179	gatgacagacaccgcgtatcagaaggagcgagacaggaagtccaacltcgcttggcatcaa	1238
	Db	1141	GATGACAGCAGCACCCGTTACGAGAGGAGCGGAGCAGCAGAGATCCAGTCCGCTGGCATCAA	1200
	OY	1239	caaacacgagtggtgccggcgcccgaagccaacacagtttgactgtggtgtgtgtccacgcg	1298
	Db	1201	CCACACCAGGAGTGCGGGGCCAAGGCCAACACAGTTCGATGTGGGCGTCTGTGGCACCCC	1260
	OY	1299	cgtgacacgcatcttccttcgttcttcactgacctgtgaccttcctccattctgcatcagtgt	1358
	Db	1261	CCTGACACCGCATTTCTCTTCTTCTTCTTCAAGCTTGCCCTTCTTCATTCTTGCCATAGGTG	1320
	OY	1359	ctccgltgtaggagt 1372	
	Db	1321	CTCCGCTGTAAGGAT 1334	
RESULT	5			
LOCUS	AC004573	85095 bp	DNA	linear INV 30-JAN-1999
DEFINITION	Drosophila melanogaster, chromosome 2L, region 21C5-21DL, P1 clone			
ACCESSION	DS07610	complete sequence.		
VERSION	AC004573			
KEYWORDS	AC004573.1	GI:4204255		
SOURCE	HTG.			
ORGANISM	fruit fly.			
	Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;			
	Pterygota; Neoptera; Endopterygota; Diptera; Brachycera;			
	Muscomorpha; Ephydroidea; Drosophilidae; Drosophila.			
REFERENCE	1 (bases 1 to 85095)			
AUTHORS	Celniker,S.E., Agbayani,A., Arcaina,T.T., Baxter,E., Blazer,R.G., Butenhoff,C., Champem,M., Chavez,C., Chew,M., Ciesiolka,L., Doyle,C.M., Farfan,D.E., Galle,R., George,R.A., Harris,N.L., Hoskins,R.A., Houston,K.A., Hummstl,S.R., Katta,K., Kearney,L., Kim,E., Lee,B., Lewis,S., Li,P., Lomolan,M.A., Mazda,P., Moshrefi,A.R., Moshrefi,M., Nixon,K., Pacleb,J.M., Park,S., Pfeiffer,B., Poon,L., Sequelair,A., Sethi,H., Snir,E., Svirskas,R.R., Wan,K.H., Weinburg,T., Zhang,R., Zieran,L.L. and Rubin,G.M.			
TITLE	Sequencing of Drosophila chromosome 2L, region 21C5-21DL			
JOURNAL	Unpublished (1998)			
REFERENCE	2 (bases 1 to 85095)			
AUTHORS	Celniker,S.E., George,R.A., Galle,R., Svirskas,R.R., Hoskins,R.A., Agbayani,A., Arcaina,T.T., Baxter,E., Blazer,R.G., Chavez,C., Chew,M., Doyle,C.M., Farfan,D.E., Flanagan,J., Houston,K.A., Humstall,S.R., Kim,K., Kearney,L., Kim,S.B., Lee,B., Lomolan,M.A., Mak,J., Mazda,P., Mok,M.S., Moshrefi,A.R., Moshrefi,M., Nixon,K., Pacleb,J.M., Park,S., Pfeiffer,B., Punch,E., Snir,E., Twomey,B., Wan,K.H., Whitelaw,K.R., Yee,A., Zhang,R., Zieran,L.L. and Kimmel,B.E.			
COMMENT	Direct Submission Submitted (11-Apr-1998) Drosophila Genome Center, Lawrence Berkeley Laboratory, MS 64-121, Berkeley, CA 94720, USA On Jan 30, 1999 this sequence version replaced gi:4164118. Sequence submitted by: Berkeley Drosophila Genome Project Lawrence Berkeley National Laboratory, MS 64-121 Berkeley, CA 94720 For further information about this sequence, including its location and relationship to other sequences, please visit our sequence archive Web site (http://www.fruitfly.org/sequence/) or send email pl library@fruitfly.berkeley.edu. Location: 80-26.			
TITLE				
JOURNAL				
FEATURES				
SOURCE	source			
	1. 85095			
	/organism="Drosophila melanogaster"			
	/strain="v2: cn bw sp"			
	/db_xref="taxon:7227"			
	/chromosome="2L"			

BASE COUNT	2289 a	19671 c	20319 g	22816 t	
Query Match	67.8%	Score 1044.2	DB 3	Length 85095	
Best Local Similarity	77.2%	Pred. No. 5.8e-171			
Matches 1518	Conservative	0	Mismatches 13	Indels 436	Gaps 3
7	caacgaggtacatccgaaacaaagcgccgtataacatgacagcactccaaagataag	66			
4746	cagcaccggtacatccgaaacaaagcgccgtataacatgacagcactccaaagataag	4805			
67	gaaccagttccggtctcgtactgactgcaactacactgctgcaaatgtaatggtctctt	126			
4806	gcaccaggtttgggggtcctgactgacacactgacgtgcaaatgtaatggtctctt	4865			
127	atggaggtgactccgctgacagcgccact	156			
4866	atggaggtgactccgctgacagcgccact	4925			
157	-----gcaaccgggagcccaagaagccaa	182			
4926	agcccgctgataagtgatctgaaatcccttcacagccacggcgaccccaagaccga	4985			
183	tgtcaagggcctgagatcgctcacaagcgcgctgtgttcaacgaactcgacagatgtga	242			
4986	tgtcaagggcctgagatcgctcacaagcgcgctgtgttcaacgaactcgacagatgtga	5045			
243	gcgcgtgttccaaggaacccccaagaagctgtccctgaaatggtgtgctacataaga	302			
5046	gcgcgtgttccaaggaacccccaagaagctgtccctgaaatggtgtgctacataaga	5105			
303	catgacgagttgaaacggaacgtacacacacacgtgctgagcgaatctc	351			
5106	catgacgagttgaaacggaacgtacacacacacgtgctgagcgaatctc	5165			
352	-----	351			
5166	ccttccttgcgcctcctctccttgcgtcgcagtgagctgagaaacacagatattttccgagttc	5225			
352	-----	351			
5226	cctcaagttcacacaactgtgatgaacattttgaaaagcgtgttatctgaactgaactgaa	5285			
352	-----	351			
5286	gtgttggtttgtttgttt	5345			
352	-----	351			
5346	gtgcatttatgataatctgataatctcaatctcaactgaacgaacgaactgctgtttttctag	5405			
352	-----	351			
5406	cgatgag	5465			
412	actttgaagttcaaggaaggtctgagaccccccaatcaacactcttcaacggggaagaagt	471			
5466	actttgaagttcaaggaaggtctgagaccccccaatcaacactcttcaacggggaagaagt	5525			
472	ggcctgagtgagcgaacccaggttgacccctcaacgaacgaatgagcactccggtggtgct	531			
5526	ggcctgagtgagcgaacccaggttgacccctcaacgaacgaatgagcactccggtggtgct	5585			
532	ccagccgtgtacacagcgtactgctgaaactcaacatgcttcaacttgccccacagacaag	591			
5586	ccagccgtgtgtacacagcgtactgctgaaactcaacatgcttcaacttgccccacagacaag	5645			
592	agctgcgaagttgaagaatcggtctctggggcctgaaagttcgtcttcggaagaagcagc	651			

Db 5646 AGCTGCAAGTTGAAGATGCGCTCCTGGGGCCGTGAAGGTGCTGCGCGAGACGGCAGC 5705
 Qy 652 gcgagagagagagtccttaccagacagacagtcagtcacccgagatggaaatcgtg 711
 Db 5706 GCGAGAGAGAGTCCCTTGACACGACGACCTGGTTCACTACCGGAGATGGGAATCGTG 5765
 Qy 712 gactcgagagcccaacttctgacagcagactacagcagcagcagcagcagcagcagc 771
 Db 5766 GACTCGGAGAGCCACTTGTGACGACGAGACTACGCGCTCATGAGATACACTCTGACG 5825
 Qy 772 gctcagagagagtcctcagacagcagcagcagcagcagcagcagcagcagcagcagc 831
 Db 5826 GCTTGAAGGCGCTCTCCATGCTACACGCGCTCATCTACACCGCGCTCTGCTGCTGTC 5885
 Qy 832 atcttgccctctcagccttctgctcctcctcctcctcctcctcctcctcctcctcctc 891
 Db 5886 ATCTTGCCCTCTACGCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 5945
 Qy 892 aacgagcctcctcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 951
 Db 5946 AACGGCCGCTCATCATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 6005
 Qy 952 gtgtgtgtccaaactcactcactgtg----- 978
 Db 6006 GTGCTGTCCAACTACTCCACTTGTGGTGTGCTGCCATTTTCATTCAGTTC 6065
 Qy 979 ----- 978
 Db 6066 AGTTCTGAACCTAGATGCTAGTACCTTAACATCCGAGANTTTACCGACTTTACCGA 6125
 Qy 979 -----gtaactctcagcagcagcagcagcagcagcagcagcagcagcagcagc 1010
 Db 6126 CCACTCAACGGGCACTTCTATATTTGCACTATCTTCTACAGCAGCAGCGCTGCTGATCT 6185
 Qy 1011 gacgctcctcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 1070
 Db 6186 GAGGCTTCCACCACTGCTGAGGTTCTAGTTCTGCTGCTGCTGCTGCTGCTGCTGCTG 6245
 Qy 1071 ggcgctcctcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 1130
 Db 6246 GCGCTCTCGGAGGCGCTGAGAAAGCTGTTGACAGGAGCTGGGACAGTGCTGCTGCT 6305
 Qy 1131 ctcggtgtcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 1190
 Db 6306 CTGCTGTTCACACACACCGGAGTGCAGAGCGGAGAGACCAAGAGATGAGAGACGA 6365
 Qy 1191 cccgctcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 1250
 Db 6366 CCCCTAGAGAGAGCGAGAGCGAGTCCAGTCCGCTGGGCAATCAACACACCGAGGT 6425
 Qy 1251 ggcgagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 1310
 Db 6426 GCCGGGCGCCAGAGCCACAGTTCGACTGAGGCGCTGCTGCGACCGCGTGGACCGAT 6485
 Qy 1311 ttcctcgttctcctcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 1370
 Db 6486 TTCTCTGCTTCTTCAAGCTGCGCTTCCATTTCTGCGCATGAGTCTCGTGTAGGG 6545
 Qy 1371 atgtctcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 1430
 Db 6546 ATGTCTGAGACTCAAGGCGACATCCCAAGCGAGTGGCAGCTCAACACTGATTTGCAAT 6605
 Qy 1431 ggcgttccatgtattatgt 1490
 Db 6606 GCGATTCATATATTTATGT 6665
 Qy 1491 tggaaataaagacactctgcagcagcagcagcagcagcagcagcagcagcagcagcagc 1537
 Db 6666 TGGATAAAGGACCTCTGCGAATGTCTGCTTACAACTAGAGAGGAA 6712

RESULT 6
 AC017491/c

LOCUS AC017491 142257 bp DNA linear HTG 09-DEC-1999
 DEFINITION Drosophila melanogaster, *** SEQUENCING IN PROGRESS ***, In ordered
 pieces.
 ACCESSION AC017491
 VERSION AC017491.1 GI:6553495
 KEYWORDS HTG; HTGS_PHASE2.
 SOURCE fruit fly.
 ORGANISM Drosophila melanogaster
 Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
 Pterygota; Neoptera; Endopterygota; Diptera; Brachycera;
 Muscomorpha; Ephydroidea; Drosophilidae; Drosophila.
 REFERENCE 1 (bases 1 to 142257)
 AUTHORS Adams, M. and Venter, J.C.
 TITLE Direct Submission
 JOURNAL Submitted (09-DEC-1999) Celera Genomics, 45 West Gude Drive,
 Rockville, MD, USA
 COMMENT This sequence was identified as CDL10210988 by the submitter.
 For more information on this record e-mail to fly@celera.com.
 * NOTE: This is a 'working draft' sequence.
 * This sequence will be replaced
 * by the finished sequence as soon as it is available and
 * the accession number will be preserved.
 FEATURES
 source
 1. 142257
 /organism="Drosophila melanogaster"
 /db_xref="taxon:7227"
 BASE COUNT 38086 a 33712 c 32464 g 37995 t
 ORIGIN
 Query Match 67.8%; Score 1044.2; DB 2; Length 142257;
 Best Local Similarity 77.2%; Pred. No. 5.9e-171;
 Matches 1518; Conservative 0; Mismatches 13; Indels 436; Gaps 3;
 Qy 7 cagcaggttacatccgaacaaagcgcgcgtgaacaaatgacagcagcagcagcagcagc 66
 Db 122926 CAGCAGCTGATCATCTGGAACAAAGCGCGGTGAGCAATGACGAGACTCCCAAGATAAG 122867
 Qy 67 gcacagcttcctcgtctcgtcagcagcagcagcagcagcagcagcagcagcagcagc 126
 Db 122866 GCACGAGTTCTCGGCTCTCGGAGCTCCACTGCTGCAAAATGATGGGATCTTCT 122807
 Qy 127 atgggagctgacattccgtgcagcagcagcagcagcagcagcagcagcagcagcagc 156
 Db 122806 ATGGGAGCTGATCTCCGTGCGGAGCGGCACTGTAACTTTACGGCTCCGAGTTGCATA 122747
 Qy 157 -----gcacgc 182
 Db 122746 AGCCGCTGATTAAGTGTATCTGAAATCCCTCCAGCCACCGCGGAGCCCAAGAGCCAA 122687
 Qy 183 tgttaagcctctgatatgcctccagcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgc 242
 Db 122686 TGTAAAGCCTCTGATGCTCCACAGCGCGGCTCTTCAAGCAATGACGACGATGTGCA 122627
 Qy 243 ggcggtgtctcagaagaccccccagcagcagcagcagcagcagcagcagcagcagcagc 302
 Db 122626 GCGGCTGTCCAAAGAACCCCAAGAGAGTGTCCCTGGAATGTGTGCACTCATAGA 122567
 Qy 303 catcagcagcttgaagcgcagcagcagcagcagcagcagcagcagcagcagcagcagc 351
 Db 122566 CATGAGACGAGTTGAAGCGGCAAGCTGACACCACTGCTGCTGATCTCGTAATCTCCAGT 122507
 Qy 352 ----- 351
 Db 122506 CCTCTCTGCT 122447
 Qy 352 ----- 351
 Db 122446 CCTCAGTTCAACAACCTGATTAAGCAATTTTGAAGAGCGTTTATCTGAAGTGAAGTGA 122387
 Qy 352 ----- 351
 Db 122386 GTGTGTGTTGTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTT 122327

Oy	352	-----	351
Dh	122326	GTGCATTTATGAATATCTGGAATATCCAAATCACTAACCAGACGAATCGTGTTTCTAG	122267
Oy	352	cgatgagagaagcggagcgctgttggaacccgtcacataatgacaacatcagcgagatc	411
Dh	122266	CGATGGAGAAGACGAGAGACCGCGTGTGGCAACCGTCACAATATGACAACTACAGCAATTC	122207
Oy	412	actttgaagttccaaagagttcttgagccccccaatgaactctcaacggcgacgaaggtc	471
Dh	122206	ACTTTGAAGTTCACGAGAGAGTGTGGAACCCCCCAATATCACTCTTCAACGGCGAGAAGGT	122147
Oy	472	gacctgatgycggaagaccagtgtaacctcaagccaaagatgycacacttcggttgatgctc	531
Dh	122146	GACCGTATGGCGCAAAACCCAGTAGTACCCTGACGACAGATGGCCACTTCGCGTGGATGCTT	122087
Oy	532	ccagcctgtgtacacggcctactctgcgaactcaaatgtcatcaacttgcgccacgaacag	591
Dh	122086	CCAGCCGCTGTACAGCGGCTCTACTGTGGAACCTCAACATGTCTCAACTGGCCCCGACACAAACGAC	122027
Oy	592	agcttcgaagttgaaagatcggctctctggggcccttgaaaggttcgtccgcgcggaagacggagc	651
Dh	122026	AGCTCAAGTTGAAGATGGCTCTCGGGGCTTGAAAGTCTGCTCTGCCGGAACGGCAGC	121967
Oy	652	gcgaaagagaagatccctctgacacacgaagacctgtgtcaagaccggaatgagaaatcgtg	711
Dh	121966	GCGAGAGAGAGATCCCTCTTGACACGACGACACTGGTTCAGTCACCCGAGAGTGGAAATGTGTG	121907
Oy	712	gaccctcgagagcccaattgtcagtcagagactacagagctataatgagatgactctctgacg	771
Dh	121906	GACTCGGACGAGCCCACTTTGTGATGATCAGATACACTAGGCTCAATAGAGATACACTCTGACG	121847
Oy	772	gctcaagcggcgctctcccatgtacacggccgtcatctacaaaccgcgctctgcatcgtc	831
Dh	121846	GCTCAGCGGCGCTCTCCATGTATACAGGGCGGTCAATACAAACCCGCGTCTGTGATGTC	121787
Oy	832	atctctgagccctctaagctcttctgtctgctctcccaatggcgcggaagaatcaatgac	891
Dh	121786	ATCTCGGCCCTCTAGGCTCTTGTGGCTGCTCCCAACATGGCGGCGGAAGATCATCATATC	121727
Oy	892	aacggcctgcacatcatcgtatgcgcgcctctctcatgtactctgcgcacgcctctcgcga	951
Dh	121726	AACGGCTGCTCATCATCAGTGTATGCGCGGCTTCCTCAATGATCTGGCCAGCTCTGTGCCA	121667
Oy	952	gtctcgtctccaaactcaactcaactgtgtg-----	978
Dh	121666	GTGCTGTCTCAACAAATATCTCCACTGTGGTGAGTGCTCCCATTTGCCATTTCAATCAGTTTC	121607
Oy	979	-----	978
Dh	121606	AGTTCTGAAACTGATGCTTAAGTACCTTAACATCGCAGATATTTACGACAACTACCGA	121547
Oy	979	-----	1010
Dh	121546	CCACTACACCGGGCAATTCCTATTAATTTGCAAGTAACTTTCTACAGCACCGCTGCTGATCT	121487
Oy	1011	gagcgtctcccaactcgttcgaagttctagtcttgttactctgacagcgagcaagaacagag	1070
Dh	121486	GAGCGTCTCCACCAATGCTCGAGGTTCTAGTCTGTACTCTTGCCCAAGCAAGCAACAAGAG	121427
Oy	1071	gcgcctgcggaagcgctcgaaaaagctgtctcaagggcaccttggcgacgtggtctgtct	1130
Dh	121426	GCGCGTCCGGAAGCGCTGNAAAAAGCTGTGTGCACAGGGCACCTGGGCACGTGGCTGCTGT	121367
Oy	1131	ctcgggtgtctcagcaaccacttggcgagtcgcgagcgcggaagaacaaagatlgagcgagca	1190
Dh	121366	CTCGGTCTTACGCAACACCGCGAGTGTGCGACGCGAAGACCAAGAGATGTGACGAGCA	121307
Oy	1191	ccgcctacaagaagcggaacggcggaagagttcgaatcgcgcttggtgacaaacacccgaggt	1250
Dh	121306	CCCGTACAGAGAGGCGGACGAGGATCTCAGTCCCTGGGCAATCAACCAACCGAGGT	121247

Qy	1251	gcggagcgccaaagccaaacgafttgcacgtgagcgatgcgtcgtcgcgcacccgcgttgaccgat	1310
Db	121246	gccggcgcccaagcccaaccagtttgcacgtgagcgatgcgtcgtcgcgcacccgcgttgaccgat	121187
Qy	1311	tctctcgatttcccttcagccttgacccctccatctctgcacatcagtgctccgtglaag	1370
Db	121186	ttcccttgatttcccttcagccttgacccctccatctctgcacatcagtgctccgtglaag	121127
Qy	1371	atgctggaagctcaaggccacatcccaagccaaigtgcaccttgacacgtttgcatt	1430
Db	121126	atgctcgagactcaagccacatcccaagccaaigtgcaccttgacacgtttgcatt	121067
Qy	1431	gcgaattcatgatttaataigtgtgctgcaactataataattaaatgaaagccttga	1490
Db	121066	gcgaattcatgatttaataigtgtgctgcaactataataattaaatgaaagccttga	121007
Qy	1491	tggaaataaagacacctgcgcgaatgtcgttcaaaaaaaaaa	1537
Db	121006	tggaaataaagacacctgcgcgaatgtcgttcaaaaaaaaaa	120960

RESULT	7
AC008002/c	
LOCUS	182726 bp DNA linear INV 06-APR-2001
DEFINITION	Drosophila melanogaster, chromosome 2L, region 21D-21E, BAC clone BACRA4BE08, complete sequence.
ACCESSION	AC008002
VERSION	AC008002.3 GI:13559545
KEYWORDS	HTG.
SOURCE	fruit fly.

ORGANISM
Drosophila melanogaster
Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
Pterygota; Neoptera; Endopterygota; Diptera; Brachycera;
Muscomorpha; Ephyroidae; Drosophilidae; Drosophila.
1 (bases 1 to 182726)
Celniker, S.E., Adams, M.D., Kronmiller, B., Tyler, D., Wan, K.H.,
AUTHORS

Holt, R.A., Evans, C.A., Gocayne, J.D., Amaratides, P.G., Brandon, R.C., Rogers, Y., An, H., Baldwin, D., Bonzon, J., Beeson, K.Y., Busam, D.A., Carlson, J.W., Center, A., Chame, M., Davenport, L.B., Dietz, S.M., Dodson, K., Dorsett, V., Doup, L.E., Doyle, C., Dresnek, D., Fartan, D., Ferreira, S., Frisoe, E., Galle, R.F., Gary, N.S., George, R.A., Gonzalez, M., Honck, J., Hoskins, R.A., Hostin, D., Howland, T.J., Ileguam, C., Jatali, M., Kruse, D., Li, P., Mattei, B., Moshrefi, A., McIntosh, T.C., Moy, M., Murphy, B., Nelson, C., Nelson, K.A., Nunoo, J., Paele, J., Paradis, V., Park, S., Patel, S., Pfeiffer, B., Phuanangvong, S., Pittman, G.S., Puri, V., Richards, S., Scheeler, F., Stapleton, M., Strong, R., Swiskas, R., Tector, C., Williams, S.M., Zaveri, J.S., Smith, H.O., Rubin, G.M. and Venter, J.C. 21E

Sequencing of *Drosophila* chromosome 2L, region 21D-21E

Unpublished

TITLE
JOURNAL

REFERENCE 2 (bases 1 to 182726)
AUTHORS Celniker,S.E., Agbayani,A., Arcaine,T.T., Baxter,E., Blazej,R.G.,
Brennerhoff,C., Chamee,M., Chavez,C., Chev,M., Cleistolis,L.,
Doyle,C.M., Farfan,D.E., Galle,R., George,R.A., Harris,N.L.,
Hoskins,R.A., Houston,K.A., Hummasti,S.R., Karra,K., Kearney,L.,
Kim,E., Lee,B., Lewis,S., Li,P., Lomotlan,M.A., Madad,P.,
Moshirefi,A.R., Moshirefi,M., Nixon,K., Pacleb,A.M., Park,S.,
Pfeiffer,B., Poon,L., Sequiera,A., Sethi,H., Snir,E.,
Svyrkas,R.R., Wan,K.H., Weinburg,T., Zhang,R., Zieran,L.L. and
Rubin,G.M.
TITLE Direct Submision
JOURNAL Submitted (09-JUL-1999) Drosophila Genome Center, Lawrence Berkeley
Laboratory, MS 64-121, Berkeley, CA 94720, USA
COMMENT On Apr 6, 2001 this sequence version replaced gi:16532012.

Berkeley Drosophila Genome Project
Lawrence Berkeley National Laboratory, MS 64-121
Berkeley, CA 94720

This sequence was assembled using end sequences from a whole genome shotgun and from subclones of this BAC and its neighboring clones. For further information about this sequence, including its location and relationship to other sequences, please visit our sequence archive web site (<http://www.fruitfly.org/sequence/>) or send email


```
/gene="CG11911"
/product="CT37076"
/db_xref="FLYBASE:Fpan0011911"
/db_xref="FLYBASE:FBgn0031249"
complement(<13648..>14580)
/gene="CG11911"
/map="21B7-21B8"
/db_xref="FLYBASE:Fpan0011911"
/db_xref="FLYBASE:FBgn0031249"
complement(13743..14576)
/gene="CG11911"
/note="CG11911 gene product"
/codon_start=1
/db_xref="FLYBASE:Fpan0011911"
/db_xref="FLYBASE:FBgn0031249"
/protein_id="AAE51508.1"
/db_xref="GI:7296216"
/translation="MKLITVTLVLAVALAAGAKLSDKLALVPSFATGVINGTENE
PHSAPYIVSLATNYLKHSHICGTLINKDWIVTAHICISEPVGMSIAGLHTRAVEE
LTQOROVDFRHKYTGKGVDPDIALHNGVIFENWOPATLPSREOVHEEHL
GKNGOPKSYTFPSGAKTQVTLTOLNREKELPEAPLAEISNCSISQOSKSN
GDSGGLVPEFTNAPSELIIVSMGYIPCLANPSTYTVSATIDWITITQSYTKL
F"
complement(join(49281..49459,49730..50200,50254..50554,
50786..51230))
/gene="CG11907"
/product="CT3136"
/db_xref="FLYBASE:Fpan0011907"
/db_xref="FLYBASE:FBgn0031250"
complement(<49281..>51230)
/gene="CG11907"
/map="21C1-21C1"
/db_xref="FLYBASE:Fpan0011907"
/db_xref="FLYBASE:FBgn0031250"
complement(join(49281..49459,49730..50200,50254..50554,
50786..51049))
/gene="CG11907"
/note="CG11907 gene product"
/codon_start=1
/db_xref="FLYBASE:Fpan0011907"
/db_xref="FLYBASE:FBgn0031250"
/protein_id="AAE51506.1"
/db_xref="GI:7296214"
/translation="MPSGSDTRRLVDALEDDDELITGEGSFVPAFRPDHHDROCLM
ERODDEVIVLSNEPSGRLETVLYFYLLIGSTMPNPFPTAEYKMYKRNASINT
DLSEELPLDKSFCDLALATISGTLFLNLNATGHNVSRTMUTLMMTLLEFV
TTGFEVINTDQEQFLILIVLNLNISAATSGALYGALEFSEFTTAVVSGA
LGGILFALFVLVLAEDPTGNTAFIFIVIGSVLLICIVCYLLARKPFRYVLEGG
DKKVIKAVSHNNGSAGEGLEPILROVMSKTYLHAISLALYTTLSVPYAVTL
MSEYGHVWTAMVMFMALNSGYFTNLLIMAPRKVRQHEKELASSIMAAALSCMAV
GSLSLIVFYOML"
join(66682..69037,75067..75450,76257..76302,76864..77132,
77259..77995)
/gene="al"
/note="Nucleotide sequence of the Celera sequence differs
from the published sequence for this transcript."
/product="CT12963"
/db_xref="FLYBASE:Fpan0003935"
/db_xref="FLYBASE:FBgn0000061"
<66682..>77995
/gene="al"
/note="CG3935"
/map="21C1-21C2"
Query Match 67.8%; Score 1044.2; DB 3; Length 302473;
Best Local Similarity 77.2%; Pred. No. 5.2e-171;
Matches 1518; Conservative 0; Mismatches 13; Indels 436; Gaps 3;
```

```
|||||
Db 235773 GCACAGTTTCCGGHCTCGAGACTGCACACTACTGCAAAATGTAATGGGATGCTT 235832
Qy 127 atgggctgaattccgtgcgaagcgcaact----- 156
Db 235833 ATGGGGCTGACTTCCTGCGACGGGCCACTGTGAAGTGTATACGGCTCCGAGTTGCATA 235892
Qy 157 -----gccacgcggagaccaccaagaagccaa 182
Db 235893 AGCGGCTATATAGTGTGATCTCGAATCCCTTCACGCCACCGCGGAGCCCAAGAACGCCAA 235952
Qy 183 lgtcaaggcccttgatcgcctccacgcgcgtgttcaagactaagacaagatgtgca 242
Db 235953 TGTCAGGCGCCTGGATGCGCTCCACAGCGCGGCTGTTCAAGAACTACGACAGATGTGCA 236012
Qy 243 gccggtgttccaaggaaccccccaagacgtgtccctgtgaatagtgtgtacactaata 302
Db 236013 GCCGGTGTCCAGGAAGAACCCCGACGACGTCCTGGAAATGGTGTGACCTACATGGA 236072
Qy 303 catcgacgagttgaacggaagctgacacccactgctggttgatctc----- 351
Db 236073 CATGACGAGTTGAACGGCAAGCTGACCCACTGCTGCTGATCTCGTAATCTCAGT 236132
Qy 352 ----- 351
Db 236133 CCTCTCCTTCGCTCTCCTCTCTGTCGAGTAGCTGAGAAACGATATTTCCAGTTC 236192
Qy 352 ----- 351
Db 236193 CCTCAAGTTCACAACTCTGATAGCAATTTGAAAGCGTGTATCTGAACTGAACTGA 236252
Qy 352 ----- 351
Db 236253 GTGTTGTTTGTGTTGTTTGTGTTTGTGTTTGTGTTTGTGTTTGTGTTTGTGTTTGTGTT 236312
Qy 352 ----- 351
Db 236313 GTGCATTTATGAAATATCGAATATCCATATCCATCAACGTAACGAAAGCAATGGTGTTCAG 236372
Qy 352 cgaatggaagaagcgaagcgtgtgcaacogtcaataatgacaatacgaagatc 411
Db 236373 CGATGGAAGACGAGAGGACGCGTGTGCAACCGTACAAATATGACAAATCATCAGCAGATC 236432
Qy 412 acttgaagtcgaagcgtgtgtgaccccccaatacaactcttaacaggggaagcgaagt 471
Db 236433 ACTTGAAGTCCAGGAGGAGTGTGACCCCAATCAACACTTAAAGCGGAGCAAGGT 236492
Qy 472 ggcctgaatgcgaagaccaggtgacccctcaagccacgaatggccaactcgtgtgagtct 531
Db 236493 GGCCTGATGCCGGAACCCAGGTACCCCTCAACCCAGATGGCCACTTCCGGTGAATGCT 236552
Qy 532 ccagcgtgtacaagcgtactatgcaactcaacatgactcgaactgcccccaagacag 591
Db 236553 CCAGCGGTGTACACGCGCTACTAGCGAATCAACATGCTCAACTGCCCCACGACACAGCAG 236612
Qy 592 agctgcaagttgaagatgctcctcgtggcctggaaggtctgtctgccccgaagacg 651
Db 236613 AGCTGCAAGTTGAAGATGCGCTCTGGGGCCCTGAAGGTCTGTCTCCGGAAGACGGCACG 236672
Qy 652 gcaagagagaatccctgacacgaacgaacgtgtgtcagtcacccggaatgggaaatcgt 711
Db 236673 GCGAAGAGAGAGTCCCTTGACACGACGACACCTCGTTCAGTCAACCGAATGGGAATCGTG 236732
Qy 712 gactcgagaccactgtgtcaagtcagaactaagcgtacatagtagtactcga 771
Db 236733 GACTCGAGGCCCACTTGTGACGTAGACTACTACGCTACATGAGATGAGTACTCTGACG 236792
Qy 772 gctcaagcggtcctctcatatgtaacgagcggtacatctacaacccggtctctgactc 831
Db 236793 GCTCAGCGGCGCTCTCATGATGACAGGCGCGTCACTCAACACCCGCGCTCTGATCTCTC 236852
Qy 832 atctcgcctctcagccttctggtcctcccaatggcgcggaagaatcatgact 891
|||||
```


LOCUS	DEFINITION	ACCESSION	VERSION	KEYWORDS	SOURCE	ORGANISM	REFERENCE	AUTHORS	TITLE	JOURNAL	FEATURES	BASE COUNT	ORIGIN
AR071400	Sequence 1 from patent US 5910582.	AR071400	AR071400.1	GI:72222288	Unknown.	Unknown.	Unclassified.	1 (bases 1 to 2068)	Ellisott,K.J., Ellis,S.B. and Harpold,M.M.	Human neuronal nicotinic acetylcholine receptor compositions and methods employing same	Patent: US 5910582;A 1 08-JUN-1999;	410 a 634 c 582 g 440 t	2 others
										Location/Qualifiers			
										1..2068			
										/organism="unknown"			

Query Match	7.1%	Score 109.4;	DB 6;	Length 2068;
Best Local Similarity	49.9%	Pred. No. 2.3e-09;		
Matches 399; Conservative	0;	Mismatches 373;	Indels 27;	Gaps 4;

QY	297	catatacatcgacgagtttgaaacgcaagctgtgacacaccacctctgctcgtgaatctccgatg	356
Db	450	CTTCATGATGGATGTGGAAAMACCAATATGTGACACCAACCGTCTGGCTTAAACAGAGATG	509
QY	357	gagagacgagagcgcgctgtgcaacgcgttcaaatatgacacgtccagcgagataactt	416
Db	510	GAGCGACTTCAAACTGCGCTGGAAACCCCGCTATTTTGGCAACATTCACATCTCTAAGGT	569
QY	417	gaagtcacgagagctctgagccccccaatcaactcttcaac---ggcgacgaagctgg	473
Db	570	CCCTTCTGAGATGATCTGGATCCCGACATTTGTTCTTCAACAAMNNTGGGAGTTTGC	629
QY	474	ccttgatgcccgaaccaccaggtgaaccttaagccagatgtgcaccttcggtgtgacctcc	533
Db	630	AGTGAACCCACATGACCCAAAGGCCACCTCTTCTCCACGGGCACTGTGCACTGGGTGTC	689
QY	534	agccgtgtacacgagctactctgaaactcaaatgtctcaactgtgcccacgaacagag	593
Db	690	GGCCATCTTCAAGAGACTCTCTGCAGATGACGTCACTTCTTCCCTTCGACACACAGAA	749
QY	594	ctgcgaagtgaagaatcgcgtctcttgggagctgaaaggtgcgtctgcgagaaacgagcagc	653
Db	750	CTGCAAGATGAAGTTGGCTCTCTGGACTTATGACAAGGCCAAGATGCG---ACCTGGAGCA	806
QY	654	gagagagagagtcctcttgacacacgacgacctggttcagtcacccgaggtggaaatcgttga	713
Db	807	GATGGAGCGAGACTGTGGAGCTTAAGGACTTACGGGAGAGCGGCGAGTGGGCTATGTCAA	866
QY	714	ctcggagagcccaacttgtacgtacgagactacgaggtacatacagag-----	759
Db	867	TGCCACGGGCACTTACAAACAGCAAGAAGTACAGCTGTGCGCGAGATCTACCCGACGT	926
QY	760	----tacaatcgaagctctcagcgcgctccctcccaatgtacaacgscgtcatctacacac	815
Db	927	CACCTTACGCGCTTGTGATCCGGCGGCTGGCGGCTTTTCAACACATCAACCTATATATCC	986
QY	816	cgcgtctcgtacatcgatcatctctgacctcgaacctctcgtcgtcgtccccaatgagcg	875
Db	987	CTGCCGTGATCTCTGCTGCTCAGTGTGCTGTCTTCACTCGCCCTCGAC---TGGGG	1043
QY	876	cgagaagatcatgatcaaacgacctgtcatcatctgtatgcgcgccttccatgtactt	935
Db	1044	CGAGAGATCAACGCGTGTGATTTGCGTGTGCTGTACACTACCGTCTTCTTCCTCTCAT	1103
QY	936	cgccacgactcccgccagagctgtgtccaacaactctcaactgtggtgaatcttacaagac	995
Db	1104	CAGTAGATCATCTCCGTCCACCTCGCTGCTCATCCCGGCTCATCGGAGTACTCTGCTTT	1163
QY	996	cagctgtcgtatcttgagcgttccacaacatcgtcaggttcaagtctcttacctgtgcaac	1055

Db	1164	CACCATGATCTTCCGTACCCCTGTCATCATCACCGTTCTGTGTCATCAATGTTGGACCA	1233
Qy	1056	aggcacagcacacagagcgcg	1074
Db	1224	CCGCTCCCCCGACCCACAC	1242

AR173184			
LOCUS	AR173184	2068 bp	DNA
DEFINITION	Sequence 1 from patent US 6303753.		linear
			PAT 17-DEC-2001

ACCESSION	ARI73184
VERSION	ARI73184.1
	GI:17912675

SOURCE	Unknown.
ORGANISM	Unknown.

Unclassified (base)

AUTHORS
Elliott,
Horton

Human the methods

JOURNAL Patent:
FEATURES

source

BASE COUNT	410
ORIGIN	

ORIGIN

Query Match

Best Local Similarity
Matches 399: Cor

Our 207 participants

27 / calaguala

450 CALICGALIGI
 DO

QY 357 gagagacgac

Db 510 GAGCGACTAC

QY 417 gaagtcagc

Db 570 CCCTTCTGAC

QY 474 cctqatqcc

630 8278A000A0

[illegible]

QY 534 agccgctgtac
111111111

Db 690 GGCCATCTAC

QY 594 ctgcaagttc

Db 750 CTGCAAGATC

QY 654 gagaggagag

Db 807 GATGGAGCAG

[illegible]

1

Db 867 TGGCACGGGC

QY 760 ----tact

Db 927 CACCTACGCC

QY 816 cgcgtcctgc

Db 987 CTGCCCTGCTC

BASE COUNT	ORIGIN	410 a	634 c	582 g	440 t	2 others

Query Match	7.1%	Score 109.4;	DB 6;	Length 2068;
Best Local Similarity	49.98;	Pred. No. 2.3e-09;		
Matches 399; Conservative	0;	Mismatches 373;	Indels 27;	Gaps 4;

QY	297	catgaacatcgaacgaglttgaaacggaagctgaccaccaccacgctgctgtaatcctccgatg	356
Db	450	CATCATGTGGATGATGAGAAACCAAAATGATGACACCAACGTCCTGGCTAAACACAGAGTG	509
OY	357	gagaacagaggaagcgctgtgtggcaacccgttaacaaatctgcaacatcaacgcagatacctt	416
Db	510	GAGGACCTACCAACTGGCGCTGGAAACCCCGCTGATTTTGGCAATCACAATCTCTAAGGT	569
OY	417	gaactccagcgaggctcttggaccccccccaatcacatcttcaao--gagacgaaggtg	473
Db	570	CCCTTCTGATGATGATCTGGATCCCGGACATTTGTTCTTACAAACAAMNNTGGGAGTTGG	629
OY	474	cctgatgycgaacccaagtgtaacctcaagcagaatgaccacttcggttgaatctcc	533
Db	630	AGTGACCCACATGTGACCAAGGCCACCTCTTCTCCACGGGCACTGTGCATCTGGTCCGCC	689
OY	534	agccgtgaacacgcgcctacatcgcgaacccaacatcgcctcaacttggccccagaaacgag	593
Db	690	GGCCATTCTACAAGAGCTCTCTGCAGCATCGACGTCACTTCTTCCCTTGACCAACAGAA	749
OY	594	ctgcgaagtctgaagctcgcgctccttggcgccgaagtcgctcttgcggaagcagcagc	653
Db	750	CTGCAAGATGAAGTTGGTCTCTGGATTTATGACAAAGGCCAAGATG--ACCTGAGAGA	806
OY	654	gagagagagatcccttaccacgaacgaacctgtgttcagtaaccggaatggaatcgtgga	713
Db	807	GATGAGCAGACTGTGGACCTGAAGGACTCTGGGAGAGCGGGAATGGGCGCATCGTCAA	866
OY	714	ctcgagagcccaacttgaagtaagagactctactacggtctacatggag-----	759
Db	867	TGCCACGGGCACTTACAAACACAAAGATGACGACTGTGGCCGAGATTTACCCCGAGCT	926
OY	760	----taacatctgaacgctcaagcgcgctctcatcagtatacgaacgcgatcatatacaccc	815
Db	927	CACCTACCCCTTGCTGATTCGGGGCGCTGGCGGCTTCTTACACCATCAACTCATATCCC	986
OY	816	cgcgtctcgtacatcgtacatcctgycacctcagcctctcgtcgtctgctccccaatgycg	875
Db	987	CTGCCTGTCTATCTTCGCTGCATATGCTGCTGTTCTTCAATGCTCCCTCCGAC--TGGGG	1043

```

QY 876 cgagaagatcatgatacaaggcctgctcatcatcgatgcgcgccttcctcatgtaact 935
    |||||
Db 1044 CGAGAGATGACGCTGTGTCATTTCGTGTCATCACTACCGCTTCCTGCTGCTCAT 1103
    |||||
QY 936 cgccagctcctgcagtgctgtccaacaalactccacttggtaatcttctaagcac 995
    |||||
Db 1104 CACTGAGATCATCCGTCACACTGCTGTCATCCGCTCATCCGCGAGTACCTGCTGT 1163
    |||||
QY 996 cagcctgctgatatgagcgcttcacacatcgtcgaagttctagtctgtacctggccac 1055
    |||||
Db 1164 CACCATGATCTTCGTCAACCTGTCATGTCATCACCCTTCGTGCTCAATGTGACCA 1223
    |||||
QY 1056 aggcaagcacaaaggcgc 1074
    |||||
Db 1224 CCGCTCCCGCCAGCACCCAC 1242
    |||||

```

Search completed: July 18, 2002, 08:54:44
 Job time: 14268 sec

XX

XX 27-SEP-2001.
PD
XX
XX 23-MAR-2001: 2001MO-US09231.
PF
XX 23-MAR-2000: 2000US-191637P.
PR 11-JUL-2000: 2000US-0614150.
XX
XX (PEKE) PE CORP NY.
PA
XX
XX Venter JC, Adams M, Li PMD, Myers EW;
PI
XX WPI; 2001-656860/75.
DR
XX
XX New isolated nucleic acid detection reagent for detecting 1000 or more
PT genes from Drosophila and for elucidating cell signalling and cell-cell
PT interactions -
PS
XX
XX Claim 1: SEQ ID NO 21424; 21bp + Sequence Listing: English.
XX
XX The invention relates to an isolated nucleic acid detection reagent
XX capable of detecting 1000 or more genes from Drosophila. The invention is
XX useful in developmental biology and in elucidating cell signalling and
XX cell-cell interactions in higher eukaryotes for the development of
XX insecticides, therapeutics and pharmaceutical drugs. The invention
XX discloses genomic DNA sequences (AB101840-AB16175), expressed DNA
XX sequences (AB101840-AB16175) and the encoded proteins
XX (AB161737-AB161752).
XX The sequence data for this patent did not form part of the printed
XX specification, but was obtained in electronic format directly from WIPO
XX at ftp.wipo.int/pub/published_pct_sequences.
XX
XX Sequence 1362 BP; 294 A; 444 C; 374 G; 250 T; 0 other;

Query Match 87.8%; Score 1350.8; DB 23; Length 1362;
Best Local Similarity 99.5%; Pred. No. 7; 7e-306;
Matches 1355; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

QY 7 caagaggtatcatccgaacaaagcgctgaacaatgaacagactcccaagataag 66
DB 1 caagaggtatcatccgaacaaagcgctgaacaatgaacagactcccaagataag 60
QY 67 gacacagttccggtctgactgacactgactgactgactgactgactgactgact 126
DB 61 gacacagttccggtctgactgacactgactgactgactgactgactgactgact 120
QY 127 atggggtgacttcgtgcaagcgccactgacacgacgacgacgacgacgacgac 186
DB 121 atggggtgacttcgtgcaagcgccactgacacgacgacgacgacgacgacgac 180
QY 187 aagggccttgatgctcccaacgacgacgacgacgacgacgacgacgacgacgac 246
DB 181 aagggccttgatgctcccaacgacgacgacgacgacgacgacgacgacgacgac 240
QY 247 gtttccaaagaaaccccaagaaagtgctcctgaaatgtgtgacactacataagac 306
DB 241 gtttccaaagaaaccccaagaaagtgctcctgaaatgtgtgacactacataagac 300
QY 307 gacagagttgaacggaacgctgacacacacacacacacacacacacacacacac 366
DB 301 gacagagttgaacggaacgctgacacacacacacacacacacacacacacacac 360
QY 367 gaggcgctgtggaacacgctacaataatgaacacacacacacacacacacacac 426
DB 361 gaggcgctgtggaacacgctacaataatgaacacacacacacacacacacacac 420
QY 427 gaggctggagaccccaacacacacacacacacacacacacacacacacacacac 486
DB 421 gaggctggagaccccaacacacacacacacacacacacacacacacacacacac 480
QY 487 acccaagtgacacctcaacacacacacacacacacacacacacacacacacacac 546
DB 481 acccaagtgacacctcaacacacacacacacacacacacacacacacacacacac 540

DB 481 acccaagtgacacctcaacacacacacacacacacacacacacacacacacacac 540
QY 547 gactactggaactcaacatgctcaactggtcccaacacacacacacacacacacacac 606
DB 541 gactactggaactcaacatgctcaactggtcccaacacacacacacacacacacacac 600
QY 607 atcggtcctggggcgctgaaggtgctcctgcaagaaagcaacgacgacgacgacgac 666
DB 601 atcggtcctggggcgctgaaggtgctcctgcaagaaagcaacgacgacgacgacgac 660
QY 667 ctgagacacagac 726
DB 661 ctgagacacagac 720
QY 727 ttgttcagtaagactactacacacacacacacacacacacacacacacacacacac 786
DB 721 ttgttcagtaagactactacacacacacacacacacacacacacacacacacacac 780
QY 787 tcaatgacacgacgac 846
DB 781 tcaatgacacgacgac 840
QY 847 gactctggtgctcctcccaacatggcgcgagagaatcaatgaacacacacacacacac 906
DB 841 gactctggtgctcctcccaacatggcgcgagagaatcaatgaacacacacacacacac 900
QY 907 atgtgtatgacgacgac 966
DB 901 atgtgtatgacgacgac 960
QY 967 actccactgtgtaactctctacacacacacacacacacacacacacacacacacacac 1026
DB 961 actccactgtgtaactctctacacacacacacacacacacacacacacacacacacac 1020
QY 1027 gtcgaagttctagtcttctacacacacacacacacacacacacacacacacacacac 1086
DB 1021 gtcgaagttctagtcttctacacacacacacacacacacacacacacacacacacac 1080
QY 1087 ctgagaaagctgtgac 1146
DB 1081 ctgagaaagctgtgac 1140
QY 1147 actgacagtcgacgacgacgacgacgacgacgacgacgacgacgacgacgacgacgac 1206
DB 1141 actgacagtcgacgacgacgacgacgacgacgacgacgacgacgacgacgacgacgac 1200
QY 1207 gacgagacagagtcacgctcgtggtacacacacacacacacacacacacacacacac 1266
DB 1201 gacgagacagagtcacgctcgtggtacacacacacacacacacacacacacacacac 1260
QY 1267 aacacagttgactgggacgctgctggtgacacacacacacacacacacacacacacac 1326
DB 1261 aacacagttgactgggacgctgctggtgacacacacacacacacacacacacacacac 1320
QY 1327 agcctgac 1368
DB 1321 agcctgac 1362

RESULT 3
ABL23316/c
ID ABL23316 standard; DNA: 3798 BP.
XX
XX ABL23316:
XX
XX 26-MAR-2002 (first entry)
XX
XX Drosophila melanogaster genomic polynucleotide SEQ ID NO 21421.
XX
XX Drosophila: developmental biology; cell signalling; insecticide;
XX pharmaceutical; gene; ds.
XX
XX Drosophila melanogaster.


```

|||||
Db 1238 CTCGGTTCAGCACACCGGCGAGTCGACAGCGAGAACCAAGAGATGACGAGCA 1179
Qy 1191 ccgtaagagagagcgagacgagagagtcagtcagtcgctggatcaaccacggagt 1250
Db 1178 CCCCTAGAGAGAGGAGACGAGCAGAGTCCAGTCCGCTGGGATCAACCCACCGAGGT 1119
Qy 1251 gccggggcccaaggccaaccagtcgagctggcgctgtggccacgcgcgtggacgcat 1310
Db 1118 GCCGGGGCCCAAGGCCAACCGATTGAGTGGGCGCTGCTGGCCACCGCCGGAGCCGAT 1059
Qy 1311 ttccctgcttccttcagcctgagccttcctcctcctcctcctcctcctcctcctcct 1370
Db 1058 TTCCTTCGTTCCCTTCAGCGCTGGCTTCCTCCTCCTCCTCCTCCTCCTCCTCCTCCT 999
Qy 1371 atgtctgagagacacagagccacatcccaagcagtcgacatctgtaactagtttcgatt 1430
Db 998 ATGCTCGAGACTCAAGGCCACATCCCAAGCCAGTGGCCACTGTAAGTTCATTTGCAATTT 939
Qy 1431 gcatctcatgataatgtagtctgagaaactaataatgataatgagacatcgtta 1490
Db 938 GCGATTTTCATGATATTTAATGCTGTGCGAACTTATTAATTTAATGATGAGACTCGTA 879
Qy 1491 tggataaaggagacccctgacgaatgctgtcttacaataaataaataa 1537
Db 878 TGGATAAAGAGACCTCTGCCGAATGCTGCTTACACTAGAACGGA 832

```

RESULT 4

ABL05370/c standard; cDNA; 7041 BP.

ABL05370;

26-MAR-2002 (first entry)

Drosophila melanogaster expressed polynucleotide SEQ ID NO 10592.

Drosophila; developmental biology; cell signalling; insecticide;

pharmaceutical; gene; ss.

Drosophila melanogaster.

WO200171042-A2.

27-SEP-2001.

23-MAR-2001; 2001WO-US09231.

23-MAR-2000; 2000US-191637P.

11-JUL-2000; 2000US-0614150.

(PEKE) PE CORP NY.

Venter JC, Adams M, Li PMD, Myers EW.

WPI: 2001-656860/75.

P-PSDB: ABB61267.

New isolated nucleic acid detection reagent for detecting 1000 or more

genes from Drosophila and for elucidating cell signalling and cell-cell

interactions -

Claim 1; SEQ ID NO 10592; 21pp + Sequence Listing; English.

The invention relates to an isolated nucleic acid detection reagent

capable of detecting 1000 or more genes from Drosophila. The invention is

useful in developmental biology and in elucidating cell signalling and

cell-cell interactions in higher eukaryotes for the development of

insecticides, therapeutics and pharmaceutical drugs. The invention

discloses genomic DNA sequences (AB16176-AB160511), expressed DNA

sequences (AB101840-AB16175) and the encoded proteins

(ABB57737-ABB72072).

CC The sequence data for this patent did not form part of the printed
 CC specification, but was obtained in electronic format directly from WIPO
 CC at ftp.wipo.int/pub/published_pct_sequences.
 XX
 SQ Sequence 7041 BP; 1753 A; 1878 C; 1797 G; 1613 T; 0 other;

Query Match 35.7%; Score 549.4; DB 23; Length 7041;
 Best Local Similarity 97.2%; Pred. NO. 1.9e-118;
 Matches 559; Conservative 0; Mismatches 16; Indels 0; Gaps 0;

```

Qy 963 caataccacactgttgtaatccttcaacagcaccagcctgtgatctgagcgtctcacc 1022
Db 6907 CATTCCTATATATTCAGATTAATCTTCAAGACAGCAGCCTGCTGATCTGAGACGCTCCAC 6848
Qy 1023 catctgcagaggttcagatctctgtaacctggtgacagagacagagagcgctgcgga 1082
Db 6847 CATCGTGGAGGTTCTAGTTCGTACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 6788
Qy 1083 ggcgtgagaaagctgtctgacagcagcagcagcagcagcagcagcagcagcagcagcag 1142
Db 6787 GGCCTGAGAAAGCTGTGTCACGGGCACTGGGCACTGGGCACTGGGCACTGGGCACTGGGCA 6728
Qy 1143 caccactgagcagtcgacagcgagagagacccaagagatgagacagcagcagcagcagcag 1202
Db 6727 CACGACGCGGAGTCCGAGCGGAGAGAACCAAGAGATGAGCAGCAGCAGCAGCAGCAGCAG 6668
Qy 1203 ggcgagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcag 1262
Db 6667 GCGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 6608
Qy 1263 ggcgaacagctgagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcag 1322
Db 6607 GGCACACAGTTCGACTGAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 6548
Qy 1323 cttaagcctgagccttcctcctcctcctcctcctcctcctcctcctcctcctcctcctcct 1382
Db 6547 CTTACGCTGCGCTTCCTCGCATTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTC 6488
Qy 1383 caagggcaacatcccaagcagtcgcaactggaactgatttgatctgagatttcagt 1442
Db 6487 CAAGGCCACATCCCAAGCCAGTGGCAGCTGCACTGCACTGCACTGCACTGCACTGCACT 6428
Qy 1443 attaatgtagtctgagaaactaataatgataatgagagcctcgatgataaagga 1502
Db 6427 ATTTAATGTGTGTGCGAATTTAATTTAATGATGAGACCTGATGAGATTAAGGA 6368
Qy 1503 cctctgcgaatgctgtcttacaataaataaataaataa 1537
Db 6367 CCTCTGCCGATGCTGCTTCACTGAAACGGA 6333

```

RESULT 5

ABL23312/c standard; DNA; 3764 BP.

ABL23312;

26-MAR-2002 (first entry)

Drosophila melanogaster genomic polynucleotide SEQ ID NO 21409.

Drosophila; developmental biology; cell signalling; insecticide;

pharmaceutical; gene; ds.

Drosophila melanogaster.

WO200171042-A2.

27-SEP-2001.

23-MAR-2001; 2001WO-US09231.

Db 630 tgcagtgacccacatgacaaagccacacctctctccacggcgactgtgcagtgtgcc 689
 Oy 531 tccagcgtgtacacagcgctactgtgaactcaactgtctaacgtgccccacgaacaa 590
 Db 650 cccggccatctacaagagctctctgcagcatcgacgtcaactctctcccttcagaccga 749
 Oy 591 gagctgcaagtgtgaacatcgctcctggggcctggaaggtctcctcgcggagaaagcgac 650
 Db 750 gaactgcagaatgaaattgtgtctctctgtgactatgacaagccaatgacg---acttga 806
 Oy 651 ggcgagagagagatgcctcttgacacgacgacctgtgtcaatcacggaagtggaaatcgt 710
 Db 807 gcaagtgtgagacgactgtggaactgaagactactgagagcgcgagtgggccatcgt 866
 Oy 711 ggactcggagagccacttgttcagtcaggaactacgtgactacgtgacatgagag----- 759
 Db 867 caatgtccagggccacttaacaacagaagaagacgactgtgcggcgagatctaccgga 926
 Oy 760 -----facactgtgaagcgtcagcgcgctccctccatctgtatcacggcgctcatcac 812
 Db 927 cgtcacctacgctctgcacatccgcggcgctgcgcgtctctctcaacatcaactatcat 986
 Oy 813 acccggtctctgcacatcgtcatcctctgacctcctcagcctctgtgctgcctccacatg 872
 Db 987 cccctgctgtcactctctcctcactgctgctgctgtgtctcttcaactgcgcctccgac---tg 1043
 Oy 873 cggcgagaagaatcatgatcaacggcgtctcatatcgttgatcgcgccctctcatgtta 932
 Db 1044 cggcgagaagaatcatcgctgtgcatctgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 1103
 Oy 933 ctgcgccagctctgcacagtgctgtccacaatatcaactgtgtgtaattcttacaag 992
 Db 1104 catcactgagatcaatcccgctccactcgtgtgtatcccgctcactacgcgagtaactgt 1163
 Oy 993 caccagcctgtgtatctgtgagcgtctccacacatcgttcgaaggtttagttctgtactgtgc 1052
 Db 1164 gtccacatgatactctgcacacctgtgtcatcgtcatcaccgtctctcgtcaatgtgca 1223
 Oy 1053 cacaagcagaagcaagaaggcgc 1074
 Db 1224 ccaacgctcccccagaccacac 1245

RESULT 7
 ID AAQ90387 standard; cDNA: 2277 BP.
 AC AAQ90387;
 DT 30-NOV-1995 (first entry)
 XX Alpha 2 subunit of a human neuronal nicotinic acetylcholine receptor.
 DE Human nNACHr; neuronal nicotinic acetylcholine receptor;
 KW neurotransmitter; ss.
 XX Homo sapiens.
 OS
 XX Key Location/Qualifiers
 FH 166..175
 FT CDS /*tag= a
 XX MO9513299-A.
 XX 18-MAY-1995.
 XX PD 08-NOV-1994; 94MO-US12859.
 XX PF 08-NOV-1993; 93US-0149503.
 XX PR 08-NOV-1993; 93US-0149503.
 PA (SALK) SALK INST BIOTECHNOLOGY IND ASSOC.
 XX Elliott KJ, Ellis SB, Harpold MM;

XX WPI: 1995-194036/25.
 DR P-PSDB; AAR73966.
 XX New human neuronal nicotinic acetylcholine receptor alpha2 subunit DNA -
 PT used to develop prods. for detection, diagnosis and therapy and for
 PT modulating activity
 PT
 PS Claim 2; Page 43-46; 54pp; English.
 XX
 CC DNA encoding the human nNACHr alpha2 subunit was isolated from a
 CC human thalamus tissue cDNA library using corresp. rat cDNA. The
 CC insert of one clone obtd. was ligated with the insert of another
 CC clone to generate a full-length alpha 2 subunit cDNA. The DNA can be
 CC used to identify function nNACHrs. Cells contg. the DNA can be used
 CC for screening to identify cpds. which modulate the activity of human
 CC nNACHrs. The human nNACHr alpha 2 subunit can be used to product
 CC antibodies which can be used in immunohistochemistry, diagnosis and
 CC therapy. The nucleic acids can be used for analysing disease states
 CC and creating animal models.
 CC
 XX Sequence 2277 BP; 451 A; 696 C; 643 G; 487 T; 0 other:
 SO

Query Match 6.9%; Score 106.8; DB 16; Length 2277;
 Best Local Similarity 49.9%; Pred. No. 3.1e-15;
 Matches 400; Conservative 0; Mismatches 372; Indels 30; Gaps 4;

Oy 297 catagacatcgagatgtgaacggaagctgacacccactgtgctgtgaatctcgaatg 356
 Db 450 catgatgtgatgagaagaaccaaagatgatgaccccaagctgtcgaacaaagagtg 509
 Oy 357 gagaagcagagagcgcggtgtgcaacccgtcacaatatgacaacatcagcagatcatt 416
 Db 510 gagcgactacaacacgtgcgtgtgaacccgcgtgatttggcaacacacatctcgaagt 569
 Oy 417 gaagtcacgagagctgtgacccccaatcacactcttcaacgycgagaaag----- 470
 Db 570 cccctctgagatgatctgtgatcccgagatgtgtcttcaacaatgacatgagtgaggat 629
 Oy 471 tggcctgtgacgaaccccaagtgaccctcaagcagaatggccactcgtgtgtagtgc 530
 Db 630 tgcagtgacccacatgacaaagccacacctctctccacggcgacatgtgcagtgtgcc 689
 Oy 531 tccagcgtgtacacagcgctactgtgaactcaactgtctaacgtgccccacgaacaa 590
 Db 650 cccggccatctacaagagctctctgcagcatcgacgtcaactctctcccttcagaccga 749
 Oy 591 gagctgcaagtgtgaacatcgctcctggggcctggaaggtctcctcgcggagaaagcgac 650
 Db 750 gaactgcagaatgaaattgtgtctctgtgactatgacaagccaatgacg---acttga 806
 Oy 651 ggcgagagagagatgcctcttgacacgacgacctgtgtcaatcacggaagtggaaatcgt 710
 Db 807 gcaagtgtgagacgactgtggaactgaagactactgagagcgcgagtgggccatcgt 866
 Oy 711 ggactcggagagccacttgttcagtcaggaactacgtgactacgtgacatgagag----- 759
 Db 867 caatgtccagggccacttaacaacagaagaagacgactgtgcggcgagatctaccgga 926
 Oy 760 -----facactgtgaagcgtcagcgcgctccctccatctgtatcacggcgctcatcac 812
 Db 927 cgtcacctacgctctgcacatccgcggcgctgcgcgtctctctcaacatcaactatcat 986
 Oy 813 acccggtctctgcacatcgtcatcctctgacctcctcagcctctgtgctcctccacatg 872
 Db 987 cccctgctgtcactctcctcactgctgctgctgtgtcttcaactgcgcctccgac---tg 1043
 Oy 873 cggcgagaagaatcatgatcaacggcgtctcatatcgttgatcgcgccctctcatgtta 932
 Db 1044 cggcgagaagaatcatcgctgtgcatctgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 1103
 Oy 933 ctgcgccagctctgcacagtgctgtccacaatatcaactgtgtgtaattcttacaag 992

```

Db      1104 caccacgagatcatccgcgcacactgcgtgcatccgcgtcatcgcgagtaacctgct 1163
Oy      993 caccacgctctgtatcttgagcgctccaccatcgaggtcttaattctgtacctgyc 1052
Db      1164 gtccacatgatctctgtcacacctgtccatgcgtacacacgctcttgctcaatgtgca 1223
Oy      1053 cacaagcagaacacagaagagcgc 1074
Db      1224 ccaccgtccccccagcaccac 1245

RESULT 8
AAT48235
ID      AAT48235 standard; DNA: 2664 BP.
AC      AAT48235;
XX
XX      09-APR-1997 (first entry)
DT
XX      Neuronal nicotinic acetylcholine receptor alpha-2 subunit DNA.
XX
XX      Neuronal nicotinic acetylcholine receptor; nAChR; neurotransmitter;
XX      ligand-gated receptor; ds.
XX
XX      Homo sapiens.
XX
XX      Key      Location/Qualifiers
XX      FH      555..2141
XX      FT      CDS
XX      FT      /*tag= a
XX      PN      WO9641876-A1.
XX
XX      PD      27-DEC-1996.
XX
XX      PF      07-JUN-1996; 96MO-US09775;
XX
XX      PR      07-JUN-1995; 950S-0484722.
XX
XX      PA      (SIBL-) SIBIA NEUROSCIENCES INC.
XX
XX      PI      Elliott KJ, Harpold MM;
XX
XX      DR      WPI; 1997-065463/06.
XX      DR      P-PSDB; AAM09021.
XX
XX      PT      Nucleic acids encoding nicotinic acetylcholine receptor sub-units -
XX      used in screening to determine the effect of drugs on the receptor
XX
XX      S      Disclosure; Page 49-52; 108pp; English.
XX
XX      CC      A DNA sequence (AAT48235) codes for the alpha-2 subunit (AAM09021) of
XX      CC      the human neuronal nicotinic acetylcholine receptor (nAChR). Host
XX      CC      cells, esp. mammalian cells or amphibian oocytes, carrying alpha-2
XX      CC      nucleic acids, opt. in combination with other alpha and/or beta
XX      CC      subunit nucleic acids (see also AAT48232-34, AAT48236-41), express
XX      CC      recombinant nAChR subunits useful for identifying cpds. that
XX      CC      modulate the activity of human nAChRs.
XX
XX      SO      Sequence 2664 BP; 518 A; 815 C; 743 G; 588 T; 0 other;

Query Match      6.9%; Score 106.8; DB 18; Length 2664;
Best Local Similarity 49.9%; Pred. No. 3.2e-15;
Matches 400; Conservative 0; Mismatches 372; Indels 30; Gaps 4;

Oy      297 catagacatcgaagcgtgaagcgaagcgtgaccacacactgctgctgaatctccgatg 356
Db      839 catcgatgtgatgagagagaacaaatgatgaccacacacgtctgctaaacagagagt 898
Oy      357 gagaagcagaagcgcgctgtggaacccgtacacatgatgacacacgtacgacgatctt 416
Db      899 gagcgactacaactgcctggaaccccgctgatttctgcaacacatcatctccaggt 958
```

```

Oy      417 gaatcagcgaagctctggaaccccccaatcacactcttcaacgcgcgaagaag----- 470
Db      959 ccctctcgagatgatctcgatctcccgacattgtctctcaacaaatgacagatggggagtt 1018
Oy      471 tggcctgataggccgaaaaccaggtgaacctcagacacagatgaccacttcggtgatgac 530
Db      1019 tgcagtgaaccacatgacaaagcccaacctctctccacagggacactgtgacatgggtgac 1078
Oy      531 tcgaagcgtgtacacggcctgactggaactcaacatgctaaatggccccaacagaaga 590
Db      1079 ccgycacatcacaagaagcctccgcagacatgaagtaacctcttcccttcgacaaga 1138
Oy      591 gaactgcaagtgtgaagaatcgcctctggggcctggaaggtctcctgcggaagaacgcac 650
Db      1139 gaactgcaagaatgaagtgttgctcctcggaacttatagacaagccaagatcg---acctgga 1195
Oy      651 ggcgagagagaagctccctgaacacgaacgcctggttcaagtaacgcgagatgggaatcgt 710
Db      1196 gcagatggagcaagactgtggaactgaagactacttggagagcgcgagtgggcatalcgt 1255
Oy      711 ggactcgagagcccaacttgtcagtaagactactaaggtacatgag----- 759
Db      1256 caatgcacagggacacttaacacagaagaagtaagactgctgcgcgagatctaccccca 1315
Oy      760 -----taactctgaagcgtcagcgcctctccatgtacacagcgctacatacac 812
Db      1316 cgtcactaagccttgtaacgcgcgcgtgcgcctctcttacaacatacactcaatc 1375
Oy      813 acccgcgtctgatactgtaacactcctgacctcagacctcttgctgctcctccacatgg 872
Db      1376 cccctgcctgctatctccgcctcactgtgctggttcttactctccctccgac--tg 1432
Oy      873 cggcgaagaagatcatatacaacggcctgcatcatcgatgatacgcgcctctctca 932
Db      1433 cgcgcgaagaatcacgctgtaacttccggtgcctgtaacactacacgctctctgctgct 1492
Oy      933 ctgcgccagctctctgccaagtgtgtccaacaataactccactgtgtgtaactctacag 992
Db      1493 catcactgagatcatccgcgcacactcgtgcgtcatccgcgtcatcgcgagtaacctgct 1552
Oy      993 caccacgctctgtatctgagcgctccacacatcgctcgaaggtctgactgtacctgac 1052
Db      1553 gtccacatgatctctgtccacctgtccatcgtacatcacccgtcttgctcaatgtgca 1612
Oy      1053 cacaagcagaacagaagagcgc 1074
Db      1613 ccaccgtccccccagcaccac 1634

RESULT 9
AAV12196
ID      AAV12196 standard; cDNA: 2363 BP.
XX
XX      AAV12196;
XX
XX      DT      14-MAY-1998 (first entry)
XX
XX      DE      Human neuronal nicotinic acetylcholine receptor alpha-4 subunit cDNA.
XX
XX      KW      Human; neuronal nicotinic acetylcholine receptor; alpha-4 subunit;
XX      KW      brain tissue; screening; NACHR; antibody; ds.
XX
XX      OS      Homo sapiens.
XX
XX      FH      Key      Location/Qualifiers
XX      FT      CDS
XX      FT      /*tag= a
XX      FT      /product= "neuronal nicotinic acetylcholine receptor
XX      PN      WO9420617-A2.
XX
```


CC A DNA sequence (AA748237) codes for the alpha-4 subunit (AAW09023) of
 CC the human neuronal nicotinic acetylcholine receptor (nAChR). Host
 CC cells, esp. mammalian cells or amphibian oocytes, carrying alpha-4
 CC nucleic acids, opt. in combination with other alpha and/or beta
 CC subunit nucleic acids (see also AA748232-36, AA748238-41), express
 CC recombinant nAChR subunits useful for identifying cpds. that
 CC modulate the activity of human nAChRs.

XX
 XX Sequence 3496 BP; 590 A; 1215 C; 1084 G; 606 T; 1 other;

Query Match

Best Local Similarity 6.8%; Score 104.6; DB 18; Length 3496;

Matches 420; Conservative 0; Mismatches 409; Indels 30; Gaps 4;

QY 297 catagacatcgaacgagttgaagcgcgaagctgacaccacgtgtgtgtgctgacatg 356
 DB 450 catgacgtgagatgagaagaacacagatgatgacacgaacgtatggtgaagcagagtg 509
 DB 357 gaagagcagagagcgctgtgtgcaaccgtlcaataatgacaatcagcagatcatt 416
 DB 510 gcaagactacaagctgcgtgcggaccagctgactatgaaatgtacacctcaccatcgcat 569
 QY 417 gaagtcacagcaggtctgtgaccccccaatacattcttaacagcgagcgaagtgacct 476
 DB 570 cccctcagagctcatctgtgagcgccgacatcgtctcttaacaacatgtacgaggact 629
 QY 477 gatgagccga-----accaggtgaccctcagccacagatgagcacttcgggtgagtgcc 530
 DB 630 cgggtgacacccacacacacgaagccacactgttccatgacggcggtgacagtgacctc 689
 QY 531 tccagcgcgtgtacacgagcctactgtcgaactcaacatgtctaaccttggtcccaacaga 590
 DB 690 cccgcacattacaagaagctcttcgcagcatcgcagcgtlcaactcttcccttcgacagca 749
 QY 591 gacgtcgaagttgaagaatcgctcctggggcccgaaagtgctcctcgcggaagagcgac 650
 DB 750 gaactgcacacatgaatctgcgtcctctgacactacgacaagcacaagatcgacctgtgaa 809
 QY 651 gggagagagagagtcctcttgacacacgacgtgttcaacatcagcagcgaggaatcgt 710
 DB 810 catgacacag---ccgctgtgacacagctgacattctggagagtgagatggtatcgt 866
 QY 711 ggaactgcgagaccacttgtcagtaaggaactacacggtactatgag-----gtg 759
 DB 867 gggcgcgtgtgagcactacacacacaggaagtagagtgctgagcagagatctaccgga 926
 QY 760 -----tacacttgaagcgtcagcggtctctctcatgtatcacggtcgatctaacac 812
 DB 927 catcactatgctctgtcatcctcgaggtgctgcgctctcttcaacatcactatcat 986
 QY 813 acccggtctcgtcatcgtacactcctgagcctctcagccttctgtgctcccccacatgg 872
 DB 987 cccctgcctgctatctctcgtcactacacgtgctgcttcttactcctccctcgaa---gtg 1043
 QY 873 cggcgaagaatcatatgaacaagcgtgtcatalcatcgtatgcgcgccttccatgta 932
 DB 1044 tggcagagaagatcacgctgtgacatctccgtgtcgtgtcgtcaccgctctcctgctgct 1103
 QY 933 ctgcgcacagctctcgcaggtgtgtccaacaatactcactcgtgttaacttctacaag 992
 DB 1104 catcaccagagatcatcccgctcacccctcaactgtcatcccatccatcagcgagatcgt 1163
 QY 993 caacagcctgtgtatctgagcgtctcacaacatcgtcgaagctgtatgtaacctgagc 1052
 DB 1164 gtctacacatgattctgtcaccctcgttccatcgtcatcaacaggtcttctgttaacgtgca 1223
 QY 1053 cacagcgaagcacaagagcgctgtccggaagcgctgagaagaactgtctgacagggacact 1112
 DB 1224 ccacgctcgcgcacagcagaacacatgtccacactgtgtacgacaggttctcctgacat 1283
 QY 1113 gggcagctgtgctgctc 1131

DB 1284 cgtgcacgctgtcctc 1302

RESULT 11

ID AAS91552 standard; cDNA; 1809 BP.

XX AAS91552;

AC AAS91552;

XX 13-FEB-2002 (first entry)

DE DNA encoding novel human diagnostic protein #27356.

XX Human; chromosome mapping; gene mapping; gene therapy; forensic;

KW food supplement; medical imaging; diagnostic; genetic disorder; ss.

XX Homo sapiens.

OS WO200175067-A2.

PN 11-OCT-2001.

PD 30-MAR-2001; 2001WO-US08631.

PF 31-MAR-2000; 2000US-0540217.

PR 23-AUG-2000; 2000US-0649167.

XX (HYSE-) HYSEQ INC.

PA Drmanac RT, Liu C, Tang YT;

PI WPI: 2001-639362/73.

DR P-PSDB: ABG27365.

XX New isolated polynucleotide and encoded polypeptides, useful in

PT diagnostics, forensics, gene mapping, identification of mutations

PT responsible for genetic disorders or other traits and to assess

PT biodiversity

XX Claim 1; SEQ ID NO 27356; 103pp; English.

XX The invention relates to isolated polynucleotide (I) and

CC polypeptide (II) sequences. (I) is useful as hybridisation probes,

CC polymerase chain reaction (PCR) primers, oligomers, and for chromosome

CC and gene mapping, and in recombinant production of (II). The

CC polynucleotides are also used in diagnostics as expressed sequence tags

CC for identifying expressed genes. (II) is useful in gene therapy techniques

CC to restore normal activity of (II) or to treat disease states involving

CC (II). (II) is useful for generating antibodies against it, detecting or

CC quantitating a polypeptide in tissue, as molecular weight markers and as

CC a food supplement. (II) and its binding partners are useful in medical

CC imaging of sites expressing (II). (I) and (II) are useful for treating

CC disorders involving aberrant protein expression or biological activity.

CC The polypeptide and polynucleotide sequences have applications in

CC diagnostics, forensics, gene mapping, identification of mutations

CC responsible for genetic disorders or other traits to assess biodiversity

CC and to produce other types of data and products dependent on DNA and

CC amino acid sequences. AAS64197-AAS94564 represent novel human

CC diagnostic coding sequences of the invention.

CC Note: The sequence data for this patent did not appear in the printed

CC specification, but was obtained in electronic format directly from WIPO

CC at ftp.wipo.int/pub/published_pcl_sequences.

XX Sequence 1809 BP; 332 A; 665 C; 481 G; 331 T; 0 other;

QY 297 catagacatcgaacgagttgaagcgcgaagctgacaccacgtgtgtgctgacatg 356
 DB 450 catgacgtgagatgagaagaacacagatgatgacacgaacgtatggtgaagcagagtg 203

Query Match 6.8%; Score 104.4; DB 23; Length 1809;
 Best Local Similarity 48.0%; Pred. No. 1e-14;
 Matches 411; Conservative 0; Mismatches 421; Indels 24; Gaps 3;

OY	357	gaggagcagagagagcgctgttgggcaacggtcaacatatgacaacatcacagcagatcattt	416
Db	204	gcaagctatacaagcttgcgtcgtggagccacagctgactatgtgaatgtcaacctccatccgat	263
OY	417	gaagttccagcgaaggttctgtgacccccccaatatcaactcttcaacggtcgaaagagtgacct	476
Db	264	ccctccgagatctcatctgtgcggcgagacatgtctctcttacaacaatgctgaagggactt	323
OY	477	gatyggccgaa-----accaggttgacctcaagcagatgtgcacattccgtgtgatatgc	530
Db	324	cggcgtaaccacacctgacccaagggcccaactgtttccatagtacggcggtgtgcaatggacctc	383
OY	511	tcacgacggttgatcacggtcgactactgtgcgaacttcaacatgtccaaacttggcccccaagaaaca	590
Db	384	ccgggcatttatacaagagctccctggcagcagcatcgaagcttcaactcttcccttcgaccagaca	443
OY	591	gagctgcagaatttgaagatcgcgtctcttggggc-----ctgaagttcgtacct	635
Db	444	gaacttgcacacatgaatctgcgtctccgtgacctatacgacaagcgcaagatcgacctgttgaa	503
OY	636	gcccggagaaacggcagcgcgagagagagatgcctcttgacccaagacgacctgttcagtaacc	695
Db	504	catgcacaagccgctgtgacccaagcttggaactcttggagagatggcgagttggtcatcgttga	563
OY	696	ggagttggaaatctgtgagctcggagcccaacttgcagttcagtaagactatacgttatat	755
Db	564	tgcctgtggcgacactacaacaccagaaagtacgaatgtcgttgcgaaatctacccggaat	623
OY	756	ggagttacactctgaacggtctacgcggtctccctccatgatacggccgtcatctacacacac	815
Db	624	caactatgaccttgtaatccggctgcggtctgcaccttcttacaacataactcatcatcc	683
OY	816	cgcgtctcgtcatcgtatcctctgtgcctctcagcctctctgtcgtcgtccccaatgagcgcg	875
Db	684	ctgccttgcatactccctgccttccacgctgtcgtctcttcaactctgcctccga---gttgg	740
OY	876	cgaaagatcatgtatacagggcctgtgcatactgcagatgcagcgcgccttccatctatctt	935
Db	741	cgaagaagatacagctgtggaatctccgttgcgtcgtcgtccacagccttccctgcgtcat	800
OY	936	cgccacagctccctgcaggtctgtccaacatactccacttgtgttaattcttataagac	995
Db	801	caccgagatcatcccgctcacactccactatgttcatcccatcctacgcgagttacctgtgt	860
OY	996	cagcctctgtatctatctgagcgtctccacacatcgtcgaagttctagtctgtacctggcac	1055
Db	861	caccatgatctctgtcacctctgtccatcgttcatcagatcgttctcgtataagcttgacaa	920
OY	1056	agggcaagacaagagggcgtcggccgagagcgctcgtgaagaagctgcgtcgaagggcagctgg	1115
Db	921	cgcgtcgcacagcagcagcaaccatctgcacacatcgtgggtacgagaggtcttcccttggaatcgt	980
OY	1116	caagtgcctgcgtcctc 1131	
Db	981	gccacgcctgtcctc 996	
RESULT 12			
AAT59528			
ID AAT59528 standard; DNA; 2082 BP.			
AAT59528;			
05-OCT-1997 (first entry)			
Alpha4 subunit of mutant nACHR (TCC 248 TTC).			
nACHR; mutation: autosomal dominant nocturnal frontal lobe epilepsy;			
ADNFLE; neuronal acetylcholine receptor, ss.			
Homo sapiens.			

FF	Key	Location/Qualifiers
FT	misc-RNA	902..1168
FT		/*tag= a
FT		/note= "ion channel"
FT	mutation	1016..1018
FT		/*tag= b
FT		/note= "TCC 248 TTC is the ADNFLE- inducing mutation (see AAU59527)."
PN	AU9656247-A.	
XX		
PD	09-JAN-1997.	
XX		
XX		
PF	28-JUN-1996;	96AU-0056247.
XX		
PR	28-JUN-1995;	95AU-0003840.
XX		
PA	(UYBO-) UNIV BONN.	
PA	(UYME) UNIV MELBOURNE.	
PPA	(WOME-) WOMEN'S & CHILDREN'S HOSPITAL.	
XX		
PI	Berkovic SF, Mulley JC, Phillips HA, Propping PJ, Scheffer IE, Steinleinok, Sutherland GR, Wallace RH;	
DR	WPI, 1997-100506/10.	
DR	P-PSDB; AAM11825.	

PT CHRNA4 gene encodes neuronal nicotinic acetylcholine receptor
PT fragment - used in diagnosis of autosomal dominant nocturnal frontal
PT lobe epilepsy
XX
PS Claim 6, Fig 13; 20pp: English.

CC The DNA sequence of the normal *nacR* is given in AAT59527.
CC The DNA sequence of the *ADNFLE*-inducing mutant is given in AAT59528.
CC Mutations at codon 248, which have the effect of replacing
CC serine by another amino acid (e.g., phenylalanine) in the sixth
CC amino acid of the transmembrane domain 2 (W2) in the alpha4
CC subunit of *nacR* have been found to be associated with *ADNFLE*.
CC The primers given in AAT59529 and AAT59530 were used in the
CC amplification of part of exon 5.

Sequence 2082 BP; 367 A; 767 C; 592 G; 356 T; 0 other;

Query Match	6.78;	Score 102.8;	DB 18;	Length 2082;
-------------	-------	--------------	--------	--------------

Matches 410; Conservative 0; Mismatches 422; Indels 24; Gaps 3;

OY	297	catgagacatcgaagattgaaacgaaagcttttaccacccaatcgtcttgatattccatg	356
Db	357	catgacgttgatgagaagaccagatgatgacacgaacgtatggtgtgaagcagagtg	456
OY	357	gagagacgagagagcgctgtgtgcaacgcgtacaaatattgacaacatcagcatcatt	416
Db	457	gcaacgaattacaacagtctgcgtctgtggaccagctgcatattagagatgacctccac	516
OY	417	gaagtcacagcgaagttcttgcgaaccccccaatatcacatttcaacgycgacgaagtg	476
Db	517	ccccctcgagctcatctctgscgcggaatcgtctcttaacaaatgctgaagcggagct	576
OY	477	gatgycgcgaa-----accacgtgtgacccctatgcacgaatgycgaacttcggtg	530
Db	577	cgcggtccaccacatgacaaagagccacactgtctccatgacgcygggtgtgacatgc	636
OY	531	tccagccgtttgacacgyccttacttgcgaactacaatatgctcaactcgtcccaagaa	590
Db	637	cccgcccatattacaagagctcctctcagatcgcagcgtcacctcttccctctgcacga	696
OY	591	gagctgcgaagttgaagatcgcgtctcttggc-----ctgaagtctgtcct	655
Db	697	gaactgcgcccatgaattcgcgtctcttgcgttaccagacgaagccaagaatgcacctg	756

[illegible]

XX	PA	(WOMEN-) WOMEN'S & CHILDREN'S HOSPITAL.
XX	PA	(UYBO-) UNIV BONN
XX	PA	(UYME) UNIV MELBOURNE.
XX	PR	28-JUN-1995; 95AU-0003840.
XX	PX	
XX	PF	28-JUN-1996; 96AU-0056247.
XX	PD	09-JAN-1997.
XX	PN	AU9656247-A.
XX	FT	
XX	FT	/note= "TCC 248 TTC is the ADNFLE-
XX	FT	inducing mutation (see AAT59528)." "
XX	FT	/tag= b
XX	FT	1016..1018
XX	FT	/note= "ion channel"
XX	FT	a
XX	FT	902..1168
XX	Key	Location/Qualifiers
XX	EH	
XX	FT	misc_RNA
XX	OS	Homo sapiens.
XX	DE	nAChR; mutation; autosomal dominant nocturnal frontal lobe epilepsy
XX	ADNFLE;	neuronal acetylcholine receptor; ss.
XX	Alpha4	subunit of normal nAChR.
XX	DT	05-OCT-1997 (first entry)
XX	AC	AAT59527;
XX	ID	AAT59527 standard; DNA; 2062 BP.
XX	RESULT	13

PI Berkovic SF, Mulley JC, Phillips RA, Propping PJ;
 P1 Scheffer IE, Steinleinok, Suterland GR, Wallace RH;
 XX WPI; 1997-100506/10.
 DR P-PSDS; AAW11824.
 XX
 XX CHRNA4 gene encodes neuronal nicotinic acetyl:choline receptor
 PT fragment - used in diagnosis of autosomal dominant nocturnal frontal
 PT lobe epilepsy
 XX
 XX
 PS Claim 5; Fig 13; 20pp: English.
 XX
 CC The DNA sequence of the normal nAChR is given in AAT59527,
 CC the DNA sequence of the ADNFLE-inducing mutant is given in AAT59528.
 CC Mutations at codon 248, which have the effect of replacing
 CC serine by another amino acid (e.g. phenylalanine) in the sixth
 CC amino acid of the transmembrane domain 2 (M2) of the alpha4
 CC subunit of nAChR have been found to be associated with ADNFLE.
 CC The primers given in AAT59529 and AAT59530 were used in the
 CC amplification of part of exon 5.
 XX
 XX Sequence 2082 BP; 367 A; 768 C; 592 G; 355 T; 0 other;

Query Match	6.6%;	Score 101.2;	DB 18;	Length 2082;
Best Local Similarity	47.8%;	Pred. No. 6.1e-14;		
Matches 409;	Conservative 0;	Mismatches 423;	Indels 24;	Gaps 3;

OY	297	catgacatcgacgagctgtgaacgcgcaagctgcacccacccactctgtgcttaactccgagt	356
Db	397	catgacgctggtatgtagaagaacaccagatgtatgacacgaacgcatgtgtgtatgaacgagagt	456
OY	357	gagagacgagagcgctgtgtgcaaccgttacaatatgtacaatcacgacgatactt	416
Db	457	gcaagacatcacaaagtctgcgtggagaccagctgtgactgtgaagatgttaacctccatccgat	516
OY	417	gaagtccaagcgaggtctcgtgaccccccacaatcacacacttccaaacgaggaagctgtgcct	476
Db	517	cccttcgcagatctcatctcgtgcgcgcggacatctgtccctctcaacaatgctgtgaacggtgactt	576
OY	477	gatgvcgca-----accacagatgaaccctccacagacagaatgagcaacttccgtgtgagtcc	530
Db	577	cgcgttcaaccacatcgaccagaagccacacccctgttctcaatgacgcggctgtgcagtgtgactcc	636
OY	531	tccaagccgtgtacacgcgctactctgcgaacttcaacatgcttcaactgtgccccaagacaaga	590
Db	637	cccgccatcttaacaagagctctctgtccagcaacgaacttcttccctcttgacacga	696
OY	591	gagcttcgaagtgtgaagatctgcgtccctctgggc-----ctbaagatgtgcct	635
Db	697	gaactgcacccatgtgaattctgcctctcgtgacctacgacaagccaaagatctgcactgtgaa	756
OY	636	gcgcgagaaacgcagcgcgagagagagatccctctgacacgaacgactgtgttcagtaacc	695
Db	757	catgcacagacgcgcgtgagaccagactgtgaactctctgtggagagatggcgaagtgtgcatcgtga	816
OY	696	gagatgggaaatcgttggaactcgcgcgcgcacccactctgtccagtcaagactactaagcttaat	755
Db	817	tgcgcgtgggcctccaacaacacacagaaatgaatgcagatgcgtgtgcgagatctaacccggaaat	876
OY	756	gagatlaactctgaacgcgtctagaacgcgcgtctctccatgatacgcgcgtatctatacacac	815
Db	877	caactatgacctgttataccgcgcgcgtctcttcttacaacatcaactcatcatccc	936
OY	816	cgctctcgtcatgttactctctgtgcctctcagctctctgtcgtcgtctccccaatgctggcg	875
Db	937	ctgcctctcatctccctgcgtctcaacgcgttgcgtctcttcaactcgtccctccga---gttggg	993
OY	876	cgaaagatcatgtataacacgcgcctgtctcatcatctgcgtacgcgcgcgccttccatctactt	935
Db	994	cgaagaaagatcaacgcgtgtgcaatctccgtgtgcgtcgtctgcacacgcgtcttccctgtctcat	1055
OY	936	cgccacagctccctgcagctgtctgtccaacaatactccacttgtgttaatcttctacagac	995

Db 1054 caccagagatcaccgctccaccactgcatccaccactcgcgcgagtaacctgtt 1113
Qy 996 cagcctgctgcatctgagcgtctccacatcgttgaggttctagtctgtaccggaac 1055
Db 1114 caccatgctctgctcaccctgctccatcgtcatcaggtctgtgctcaacgtccacca 1173
Qy 1056 aggaagacacaaagagcgccgctgcggaagcgtgagaagctgtgcacgggacactggg 1115
Db 1174 ccgctgcacacgacgcacgcacaccatgcccactggtgtaacgaggtcttctcgtgacatcgt 1233
Qy 1116 cagctgctgctgtctc 1131
Db 1234 gccacgctgctctc 1249

RESULT 14

AAQ14288
ID AAQ14288 standard; DNA; 1521 BP.

AAQ14288:

16-JAN-1992 (first entry)

Human neuronal nicotinic acetylcholine receptor beta 2 subunit.

hnnachr; ss.

W09115602-A.

17-OCT-1991.

03-APR-1991; 91WO-US02311.

03-APR-1990; 90US-0504455.

(SALK) SALK INST BIOTECHN.

Harpold MM, Ellis SB, Brust P, Akong M, Velicelbi G:

WPI; 1991-325241/44.

Human neuronal nicotinic acetylcholine receptor sub-units and
DNA - used for screening for agonists or antagonists for the
receptors.

Claim 5; Fig 9; 61pp; English.

The DNA sequence encodes the beta 2 subunit of the human neuronal
nicotinic acetylcholine receptor (hnnachr). The sequence was
isolated by probing various human neuronal tissue cDNA libraries
e.g. pre-frontal cortex or brain stem cDNA with analogous rat
neuronal acetylcholine receptor subunit fragments as probes. The
sequence shows 87% homology with the rat beta 2 sequence. The
receptor subunit gene and encoded protein can be used for drug
design and screening for obtaining drugs which have agonist or
antagonist activity on NMACH receptors. See also Q114284-AAQ14287.

Sequence 1521 BP; 273 A; 499 C; 425 G; 324 T; 0 other;

Query Match 6.4%; Score 98; DB 12; Length 1521;

Best Local Similarity 48.1%; Pred. No. 3,1e-13;

Matches 414; Conservative 0; Mismatches 425; Indels 21; Gaps 4;

Qy 287 tggtaacatagacatgacagcaggttgaaagcgaagctgaccacacactgtgctga 346
Db 203 tggccacagctacatgctgcatgagcgagagatcatgacacacaaatgtctgctga 262
Qy 347 atctccagatgagagagagcgctgtgcaacgctcacatataacacacagc 406
Db 263 cccagagatgagagatcatgctcaccctggaagcctggaaggttgacacacagaga 322

Qy 407 agatcacttgaagtcacagcaggtctgagaccoccccaataacacacttcaacgcgagc 466
Db 323 aagtcgctccctccatcaaacacatcgtgctcccaagatgtgtctgttaacaacatgtc 382
Qy 467 aaggt-----ggtctgtagcgcgaacaccaggtgaccccaagcagaatggtccatcc 520
Db 383 acgcatgtacgaggtgtctctctctatccaaatgcccgtgtctccatgtagcagacatc 442
Qy 521 ggtggtctcccaacgctgtgtaacagcctactgtgaaactcaacatgtcaactgtgccc 580
Db 443 tctggtgctgcgcctgccaatccataacagagcgtatgaaatgaaagacatctccat 502
Qy 581 acgaacagacagctgcaagttgaaagatcgtctccgtggtgcgtgaagctgtccctgcg 640
Db 503 ttgaccagacagacatgacacatgaaatgctgtgtgtaactgaacccgacagagatcgt 562
Qy 641 agaacgcaagcgagagagagatcccttgtaacacagcagcactgtgtcaagtaaccgagat 700
Db 563 actgtgtgtaagaggtgagtgagc---agcctgagacgacttcaacactagtgtagt 619
Qy 701 gggaatcgtgtgactcgcgagccacttgtcagtaagcagactac-----tagagct 751
Db 620 gggaatcgtgtgactcgcgagccacttgtcagtaagcagactac-----tagagct 751
Qy 752 acatggaatcactctgacgctcagcgctcctccatgtatgacagcgctcatcaca 811
Db 680 acatcagctagactctcatctcacttcgcgaacgctcttcttaccacatcaacctcaca 739
Qy 812 caccgcgtctgcatcgtcatcctcgtccctcagcctcctgctgctcccaacatg 871
Db 740 tccctgtgtgtcatcaccctcgtcagcactcgtctctctactacgtcagac---t 796
Qy 872 gcgaggaagaaatcatgtaacagcgctgctcatcatatgtatgcgccttccatcagt 931
Db 797 gtgaggaagaaatgagctgtgcatcagtgctgtgtgcgtcagctcagctcgtcgtc 856
Qy 932 actgcgcagctcctcgtcagtgctgccaacatcaccacttgtgtaactctctaca 991
Db 857 tcatctccaagatcgtgtcctccacactcctcgaagtgctgctgagtaagtaacctca 916
Qy 992 gaaccagctgtatctatgagcgtctccacacatcgtcgaagttctagtctgtaacctg 1051
Db 917 tgtacacatgtgtgttgaaccttccatcgtacacagcgtgtgctgtcacaagctgc 976
Qy 1052 ccacaggaacacacagagcgctgcccggagcgctgagaagctgtcgaagcgacc 1111
Db 977 accacgctgcgcac 1036
Qy 1112 tgggacgtgtgctgtc 1131
Db 1037 agtgcgcgcgtgtcttc 1056

RESULT 15

AAAS15343
ID AAAS15343 standard; DNA; 1509 BP.

AAAS15343:

16-JAN-2002 (first entry)

Human CHRNA2 coding sequence.

Human; cholinergic receptor, nicotinic, beta polypeptide 2; neuronal;
CHRNA2; memory disorder; Alzheimer's disease; epilepsy; learning;
chromosome 1q21; schizophrenia; attention deficit/hyperactivity disorder;
ADHD; autosomal dominant nocturnal frontal lobe epilepsy; ADNFLE; ds.

Homo sapiens.

Key Location/Qualifiers
CDS 1..1509
FT /*tag= a

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: July 18, 2002, 10:24:32 ; Search time 2887.37 Seconds

(without alignments)
1154.074 Million cell updates/sec

Title: US-09-732-680A-1

Perfect score: 1539

Sequence: 1 attcgacacgaggtacatc.....cttacaaaaaaaaaaaaa 1539

Scoring table:

OLIGO_NUC
Gapop 60.0 , Gapext 60.0

Searched: 1797656 seqs, 10463268293 residues

ord size : 0

Total number of hits satisfying chosen parameters: 3595312

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Listing first 1000 summaries

Database :

GenEmbl: *
1: gb_ba: *
2: gb_htg: *
3: gb_in: *
4: gb_om: *
5: gb_ov: *
6: gb_pat: *
7: gb_ph: *
8: gb_pl: *
9: gb_pr: *
10: gb_ro: *
11: gb_sts: *
12: gb_sy: *
13: gb_un: *
14: gb_vi: *
15: em_ba: *
16: em_fun: *
17: em_hum: *
18: em_in: *
19: em_mu: *
20: em_om: *
21: em_or: *
22: em_ov: *
23: em_pat: *
24: em_ph: *
25: em_pl: *
26: em_ro: *
27: em_sts: *
28: em_un: *
29: em_vi: *
30: em_htg_hum: *
31: em_htg_inv: *
32: em_htg_other: *
33: em_htgo_inv: *

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Query Match Length	ID	Description
------------	--------------------	----	-------------

1	1539	100.0	1539	6	AX113259	AX113259 Sequence
2	1421	79.3	1539	2	AY005148	AY005148 Drosophila
3	1181	76.7	1334	3	DME318761	AJ138761 Drosophila
4	628	40.8	85095	3	AC004573	AC004573 Drosophila
5	628	40.8	142257	3	AC017491	AC017491 Drosophila
6	628	40.8	182726	3	AC008002	AC008002 Drosophila
7	628	40.8	302473	3	AE003589	AE003589 Drosophila
8	507	32.9	1646	3	AY061622	AY061622 Drosophila
9	144	9.4	299	11	DM159B9T	231885 D. melanoga
10	24	1.6	74135	2	AC101533	AC101533 Mus muscu
11	24	1.6	77575	2	AC096600	AC096600 Rattus no
12	24	1.6	161361	2	AC090939	AC090939 Homo sapi
13	24	1.6	178622	2	AC026020	AC026020 Homo sapi
14	24	1.6	191494	2	AL591488	AL591488 Homo sapi
15	24	1.6	239135	2	AL591607	AL591607 Mus muscu
16	23	1.5	20149	2	HS129H7B	Z54146 Human DNA s
17	23	1.5	62537	2	AC101100	AC101100 Mus muscu
18	23	1.5	110000	2	AL136360-2	Continuation (3 of
19	23	1.5	146407	2	AC004953	AC004953 Homo sapi
20	23	1.5	154998	2	AC109787	AC109787 Bos tauru
21	23	1.5	173512	2	AC068565	AC068565 Homo sapi
22	23	1.5	186026	2	AC099785	AC099785 Homo sapi
23	23	1.5	195773	2	AC073134	AC073134 Homo sapi
24	23	1.5	200622	2	AL158068	AL158068 Homo sapi
25	23	1.5	202375	2	AC109798	AC109798 Bos tauru
26	22	1.4	1135	8	AF134302	AF134302 Arabidops
27	22	1.4	1488	8	AY070424	AY070424 Arabidops
28	22	1.4	133260	2	AC024523	AC024523 Homo sapi
29	22	1.4	143769	2	AC004914	AC004914 Homo sapi
30	22	1.4	154673	2	AC079472	AC079472 Mus muscu
31	22	1.4	168463	2	AC021463	AC021463 Homo sapi
32	22	1.4	180844	10	AL606536	AL606536 Mouse DNA
33	22	1.4	187746	2	AC110003	AC110003 Homo sapi
34	22	1.4	189781	2	AC027506	AC027506 Homo sapi
35	22	1.4	217831	2	AC108162	AC108162 Homo sapi
36	22	1.4	230254	2	AC074163	AC074163 Mus muscu
37	22	1.4	239057	2	AF001578	AF001578 Homo sapi
38	22	1.4	254278	2	AC023836	AC023836 Mus muscu
39	22	1.4	293043	2	AC073707	AC073707 Mus muscu
40	21	1.4	581	9	AF029670	AF029670 Homo sapi
41	21	1.4	1632	9	AF267864	AF267864 Homo sapi
42	21	1.4	1731	10	AY012159	AY012159 Mus muscu
43	21	1.4	2030	3	AY051926	AY051926 Drosophila
44	21	1.4	2688	9	HSMB00644	AL096749 Homo sapi
45	21	1.4	6656	6	AX281211	AX281211 Sequence
46	21	1.4	6656	6	AX356484	AX356484 Sequence
47	21	1.4	23777	2	AC106371	AC106371 Rattus no
48	21	1.4	41887	6	AX059526	AX059526 Sequence
49	21	1.4	57506	2	AC109246	AC109246 Mus muscu
50	21	1.4	68759	9	AL136111	AL136111 Human DNA
51	21	1.4	88848	8	T24H24	AF075598 Arabidops
52	21	1.4	90248	8	AC026480	AC026480 Arabidops
53	21	1.4	94536	9	AC010587	AC010587 Homo sapi
54	21	1.4	117621	9	AC003950	AC003950 Homo sapi
55	21	1.4	124456	2	AC017830	AC017830 Drosophila
56	21	1.4	136037	9	AC004104	AC004104 Homo sapi
57	21	1.4	136723	9	AC074000	AC074000 Homo sapi
58	21	1.4	139771	9	AC007425	AC007425 Homo sapi
59	21	1.4	141048	2	AC007644	AC007644 Homo sapi
60	21	1.4	143720	10	AC087332	AC087332 Mus muscu
61	21	1.4	148927	2	AC087127	AC087127 Mus muscu
62	21	1.4	152558	2	AC013487	AC013487 Homo sapi
63	21	1.4	154998	2	AC109787	AC109787 Bos tauru
64	21	1.4	158698	2	AC018625	AC018625 Homo sapi
65	21	1.4	161798	2	AC068062	AC068062 Homo sapi
66	21	1.4	164773	2	AC091543	AC091543 Felis cat
67	21	1.4	167912	3	AC007650	AC007650 Drosophila
68	21	1.4	171087	3	AC0073584	AC0073584 Homo sapi
69	21	1.4	173352	2	AC015543	AC015543 Homo sapi
70	21	1.4	173475	2	AC021864	AC021864 Homo sapi
71	21	1.4	175064	9	AC093848	AC093848 Homo sapi
72	21	1.4	175384	9	AL451103	AL451103 Human DNA
73	21	1.4	177298	2	AC068751	AC068751 Homo sapi

74	21	1.4	177673	2	AL365182	147	20	1.3	2596	5	DRE315610	AJ315610	Danio rer
C 75	21	1.4	178982	2	AC105656	148	20	1.3	2636	9	AK021422	AK021422	Homo sapi
76	21	1.4	179454	3	AC007646	149	20	1.3	2680	9	BC001336	BC001336	Homo sapi
C 77	21	1.4	188701	2	AC025242	150	20	1.3	3090	3	AY051596	AY051596	Drosophila
C 78	21	1.4	190664	2	AC090984	151	20	1.3	3535	8	YSPCTL	YSPCTL	Drosophila
C 79	21	1.4	190848	2	AC097310	152	20	1.3	4415	6	AX281491	AX281491	Yeast gene
C 80	21	1.4	180222	8	ATCHRIV1	153	20	1.3	4415	6	AX346808	AX346808	Sequence
C 81	21	1.4	203291	2	AC020998	154	20	1.3	4415	6	AX348855	AX348855	Sequence
C 82	21	1.4	208418	3	AE003701	155	20	1.3	4498	10	MMU11494	MMU11494	Mus muscu
C 83	21	1.4	209003	2	AC073724	156	20	1.3	4955	9	AB036063	AB036063	Homo sapi
C 84	21	1.4	209685	2	AC026231	157	20	1.3	5068	9	HSB010925	HSB010925	Homo sapi
C 85	21	1.4	216877	2	AC105707	158	20	1.3	5236	6	AX345252	AX345252	Sequence
C 86	21	1.4	220491	2	AL445206	159	20	1.3	5340	6	AX345137	AX345137	Sequence
C 87	21	1.4	325417	2	AC079422	160	20	1.3	5817	9	AB028975	AB028975	Homo sapi
C 88	21	1.4	322436	2	AC074169	161	20	1.3	5825	6	AX323670	AX323670	Sequence
C 89	20	1.3	290	6	AX256044	162	20	1.3	5947	9	AF006513	AF006513	Homo sapi
C 90	20	1.3	290	6	AX198718	163	20	1.3	5964	6	AX344763	AX344763	Sequence
C 91	20	1.3	290	6	AX209250	164	20	1.3	6208	6	AX345096	AX345096	Sequence
C 92	20	1.3	290	6	AX209660	165	20	1.3	6352	6	AX345807	AX345807	Sequence
C 93	20	1.3	364	11	G49426	166	20	1.3	6361	6	AX251331	AX251331	Sequence
C 94	20	1.3	403	6	AX072066	167	20	1.3	6361	6	AX346736	AX346736	Sequence
C 95	20	1.3	420	6	ROEST042	168	20	1.3	6478	6	AX277961	AX277961	Sequence
C 96	20	1.3	502	6	AX255948	169	20	1.3	6478	6	AX323656	AX323656	Sequence
C 97	20	1.3	564	8	AF279895	170	20	1.3	6478	6	AX344793	AX344793	Sequence
C 98	20	1.3	590	8	SPISBIA	171	20	1.3	6608	6	AX346937	AX346937	Sequence
C 99	20	1.3	792	9	BC005905	172	20	1.3	6626	6	AX251568	AX251568	Sequence
C 100	20	1.3	797	8	AF244121	173	20	1.3	6716	6	AX346685	AX346685	Sequence
C 101	20	1.3	806	8	AF059294	174	20	1.3	6949	9	HSB297557	HSB297557	Sequence
C 102	20	1.3	892	8	AK000610	175	20	1.3	8129	6	AX251521	AX251521	Sequence
C 103	20	1.3	914	8	AY050945	176	20	1.3	8136	6	AX345457	AX345457	Sequence
C 104	20	1.3	917	10	BC009166	177	20	1.3	8136	6	AX348343	AX348343	Sequence
C 105	20	1.3	985	10	BC015940	178	20	1.3	8547	6	AX251816	AX251816	Sequence
C 106	20	1.3	1018	3	PFU41269	179	20	1.3	8547	6	AX344200	AX344200	Sequence
C 107	20	1.3	1040	10	BC006787	180	20	1.3	8547	6	AX348603	AX348603	Sequence
C 108	20	1.3	1138	9	BC000863	181	20	1.3	8801	6	AX277981	AX277981	Sequence
C 109	20	1.3	1194	9	HSU94317	182	20	1.3	8801	6	AX323672	AX323672	Sequence
C 110	20	1.3	1201	9	EC005398	183	20	1.3	8801	6	AX346643	AX346643	Sequence
C 111	20	1.3	1203	9	BC004240	184	20	1.3	9106	6	AX281327	AX281327	Sequence
C 112	20	1.3	1234	9	HSB22580	185	20	1.3	9838	6	AX281304	AX281304	Sequence
C 113	20	1.3	1241	10	AF146524	186	20	1.3	9838	6	AX345481	AX345481	Sequence
C 114	20	1.3	1242	10	MMU250491	187	20	1.3	10279	6	AX279995	AX279995	Sequence
C 115	20	1.3	1247	8	AY056230	188	20	1.3	10279	6	AX346492	AX346492	Sequence
C 116	20	1.3	1254	8	AY061131	189	20	1.3	10279	6	AX356451	AX356451	Sequence
C 117	20	1.3	1271	8	AY040022	190	20	1.3	10816	3	DME18278	DME18278	Sequence
C 118	20	1.3	1337	8	AF274589	191	20	1.3	10855	6	AX346390	AX346390	Sequence
C 119	20	1.3	1351	9	BC007001	192	20	1.3	12359	6	AX346948	AX346948	Sequence
C 120	20	1.3	1401	3	AF338731	193	20	1.3	12733	6	AX344993	AX344993	Sequence
C 121	20	1.3	1440	9	BC001442	194	20	1.3	12927	2	AC107283	AC107283	Rattus no
C 122	20	1.3	1449	10	BC021947	195	20	1.3	14006	6	AX346860	AX346860	Sequence
C 123	20	1.3	1494	9	S7450453	196	20	1.3	15782	6	AX323521	AX323521	Sequence
C 124	20	1.3	1499	8	AY050825	197	20	1.3	15881	6	AX251764	AX251764	Sequence
C 125	20	1.3	1541	8	BC018788	198	20	1.3	15881	6	AX344178	AX344178	Sequence
C 126	20	1.3	1572	8	AF275752	199	20	1.3	15881	6	AX345162	AX345162	Sequence
C 127	20	1.3	1604	8	AF370210	200	20	1.3	15881	6	AX348575	AX348575	Sequence
C 128	20	1.3	1619	3	AY070934	201	20	1.3	16173	6	AX281358	AX281358	Sequence
C 129	20	1.3	1691	10	BC002192	202	20	1.3	17251	10	AB015623	AB015623	Mus muscu
C 130	20	1.3	1719	8	SPU16143	203	20	1.3	19946	9	AL583784	AL583784	Human DNA
C 131	20	1.3	1747	8	MCMDH1A	204	20	1.3	26203	2	PFMAL13P7	PFMAL13P7	Human DNA
C 132	20	1.3	1790	5	AF071570	205	20	1.3	27746	9	HS64E4	HS64E4	Human DNA
C 133	20	1.3	1878	9	AK021503	206	20	1.3	29776	2	AC017410	AC017410	Drosophila
C 134	20	1.3	1897	9	AK054572	207	20	1.3	36030	8	F16L1	F16L1	Sequence
C 135	20	1.3	1926	9	AK025140	208	20	1.3	36034	8	AP001296	AP001296	Homo sapi
C 136	20	1.3	1943	3	DDV14	209	20	1.3	36252	9	AP001056	AP001056	Homo sapi
C 137	20	1.3	1957	9	HSB420507	210	20	1.3	36430	8	AC010506	AC010506	Homo sapi
C 138	20	1.3	1958	9	AK026674	211	20	1.3	37228	8	SPC5757	SPC5757	Homo sapi
C 139	20	1.3	2113	10	AF418207	212	20	1.3	38237	2	AC013904	AC013904	Drosophila
C 140	20	1.3	2173	6	HSU28833	213	20	1.3	38478	2	AC019893	AC019893	Drosophila
C 141	20	1.3	2174	6	AK034241	214	20	1.3	41525	2	AC110343	AC110343	Rattus no
C 142	20	1.3	2227	9	HSU85266	215	20	1.3	43089	2	AC017875	AC017875	Drosophila
C 143	20	1.3	2284	9	HSU85265	216	20	1.3	43919	9	AL446023	AL446023	Human DNA
C 144	20	1.3	2346	3	AF243046	217	20	1.3	43983	9	AL136970	AL136970	Human DNA
C 145	20	1.3	2346	3	HSU85267	218	20	1.3	44623	9	AL138964	AL138964	Human DNA
C 146	20	1.3	2581	3	DDIGP138A	219	20	1.3	47030	9	AC092310	AC092310	Homo sapi

220	1.3	47090	9	AC093233	Homo sapi	293	1.3	110000	2	AL35978_1	Continuation (2 of
221	1.3	48282	2	AC024228	Atrosidops	294	1.3	11051	9	AL353806	AL353806 Human DNA
222	1.3	49491	2	AC013847	AC013847 Drosophil	295	1.3	112401	9	AL606462	AL606462 Human DNA
223	1.3	56631	9	HS702M17	AL034427 Human DNA	296	1.3	112573	9	AL353151	AL353151 Human DNA
224	1.3	59841	2	AC102369	AC102369 Mus muscu	297	1.3	113253	2	AC092356	AC092356 Human DNA
225	1.3	61159	2	AC021911	AC021911 Homo sapi	298	1.3	113335	9	AL596268	AL596268 Human DNA
226	1.3	61278	2	AC024308	AC024308 Homo sapi	299	1.3	113515	6	AX347076	AX347076 Sequence
227	1.3	61514	2	AC087697	AC087697 Homo sapi	300	1.3	114016	9	AL365396	AL365396 Human DNA
228	1.3	62497	2	AC103865	AC103865 Homo sapi	301	1.3	114604	9	AL391376	AL391376 Human DNA
229	1.3	63859	2	AC064794	AC064794 Homo sapi	302	1.3	116215	9	AC004130	AC004130 Homo sapi
230	1.3	64737	2	AC104382	AC104382 Homo sapi	303	1.3	117613	9	AC092784	AC092784 Homo sapi
231	1.3	64737	2	AC104382	AC104382 Homo sapi	304	1.3	121101	2	AC010931	AC010931 Homo sapi
232	1.3	66206	2	AC023324	AC023324 Homo sapi	305	1.3	123391	9	HS7902P15	HS7902P15 Human DNA
233	1.3	66259	9	AL390070	AL390070 Human DNA	306	1.3	125351	2	AC104642	AC104642 Typanoso
234	1.3	66352	2	AC110047	AC110047 Homo sapi	307	1.3	127370	9	AL451044	AL451044 Human DNA
235	1.3	67780	2	AC087275	AC087275 Homo sapi	308	1.3	127462	8	F5D14	AC007767 Sequence
236	1.3	68851	2	AC101183	AC101183 Mus muscu	309	1.3	128143	10	AC003063	AC003063 Mus muscu
237	1.3	68959	9	HS735P11	AL049783 Human DNA	310	1.3	128328	9	AC005839	AC005839 Homo sapi
238	1.3	68997	2	AC080153	AC080153 Homo sapi	311	1.3	128335	9	AC093864	AC093864 Homo sapi
239	1.3	70416	9	AC095039	AC095039 Homo sapi	312	1.3	128703	9	AF207550	AF207550 Homo sapi
240	1.3	70800	2	AC099950	AC099950 Mus muscu	313	1.3	129046	9	AC008835	AC008835 Homo sapi
241	1.3	71092	2	AL590732	AL590732 Human DNA	314	1.3	129129	9	AL513324	AL513324 Human DNA
242	1.3	71142	2	AC016305	AC016305 Homo sapi	315	1.3	132122	9	AC096582	AC096582 Homo sapi
243	1.3	73778	6	AX344562	AC004562 Sequence	316	1.3	132493	9	AC004583	AC004583 Human Chr
244	1.3	75609	9	AC006456	AC006456 Homo sapi	317	1.3	132809	10	AC087262	AC087262 Rattus no
245	1.3	76618	3	AC004657	AC004657 Drosophil	318	1.3	133368	9	AL627255	AL627255 Dantio rer
246	1.3	76808	2	AC014187	AL444198 Human DNA	319	1.3	134156	9	AL357083	AL357083 Human DNA
247	1.3	78404	9	AL445198	AC014363 Drosophil	320	1.3	134210	9	AC005052	AC005052 Homo sapi
248	1.3	79432	2	AC0014363	AC006252 Homo sapi	321	1.3	134278	9	AP001822	AP001822 Homo sapi
249	1.3	82098	9	AC006252	AL359980 Human DNA	322	1.3	134376	2	AC0073440	AC0073440 Homo sapi
250	1.3	83218	9	AL359980	AC106849 Homo sapi	323	1.3	135053	2	AC094957	AC094957 Rattus no
251	1.3	83547	2	AC106849	AL589984 Human DNA	324	1.3	135240	9	AL136362	AL136362 Human DNA
252	1.3	84388	9	AL589984	AC096145 Rattus no	325	1.3	136862	9	AC000507	AC000507 Homo sapi
253	1.3	86110	2	AC096145	AP001426 Homo sapi	326	1.3	137184	9	AL353692	AL353692 Human DNA
254	1.3	87402	9	AP001426	AL645810 Dantio rer	327	1.3	137508	2	AC106622	AC106622 Rattus no
255	1.3	88118	2	AL645810	AC108387 Pan trogl	328	1.3	137515	2	DMBR12B2	AL122029 Drosophil
256	1.3	89017	2	AC108387	AL121909 Human DNA	329	1.3	139594	2	AC025085	AC025085 Homo sapi
257	1.3	89251	9	HS0812P3	AC094673 Rattus no	330	1.3	139773	2	AC067806	AC067806 Homo sapi
258	1.3	89369	2	AC094673	AC064829 Homo sapi	331	1.3	140073	9	CNS01DPO	AL133819 Human chr
259	1.3	89684	9	AC064829	AC090287 Homo sapi	332	1.3	140999	9	AL356415	AL356415 Human DNA
260	1.3	90343	9	AC090287	AL031178 Human DNA	333	1.3	141547	2	AC098660	AC098660 Rattus no
261	1.3	92558	9	HS341E18	U63630 Homo sapien	334	1.3	141794	2	AC093455	AC093455 Homo sapi
262	1.3	93443	9	HS424E5	AP000246 Homo sapi	335	1.3	143427	9	AC002432	AC002432 Human BAC
263	1.3	93805	9	HS424E5	AP000246 Homo sapi	336	1.3	143618	9	AL137856	AL137856 Human DNA
264	1.3	94212	9	AL360083	AL360083 Human DNA	337	1.3	145028	2	AC073875	AC073875 Homo sapi
265	1.3	94963	9	AL360083	AC094668 Rattus no	338	1.3	145189	2	AC021968	AC021968 Homo sapi
266	1.3	95911	2	AC094668	AC103047 Rattus no	339	1.3	145366	2	AC024945	AC024945 Homo sapi
267	1.3	96268	2	AC103047	AC009496 Homo sapi	340	1.3	145380	2	AL355542	AL355542 Homo sapi
268	1.3	96843	9	AC009496	AL161916 Human DNA	341	1.3	146565	2	AC023320	AC023320 Homo sapi
269	1.3	99249	9	AP000021	AP000021 Homo sapi	342	1.3	146814	2	AC093977	AC093977 Rattus no
270	1.3	100000	9	AP000021	AP000070 Homo sapi	343	1.3	147465	2	AC067940	AC067940 Homo sapi
271	1.3	100000	9	AP000070	AP000130 Homo sapi	344	1.3	148257	9	AC018258	AC018258 Drosophil
272	1.3	100000	9	AP000130	AP000163 Homo sapi	345	1.3	148418	5	AL590150	AL590150 Zebrafish
273	1.3	100000	9	AP000163	AP000208 Homo sapi	346	1.3	148777	2	AC023101	AC023101 Homo sapi
274	1.3	100000	2	AP000208	AC106551 Rattus no	347	1.3	148847	2	AC016039	AC016039 Homo sapi
275	1.3	102343	2	AC106551	AC108086 Homo sapi	348	1.3	149322	2	AL390880	AL390880 Homo sapi
276	1.3	104762	2	HS67A5	AL035594 Human DNA	349	1.3	149837	2	AC015504	AC015504 Homo sapi
277	1.3	105100	9	HS67A5	AC084013 Homo sapi	350	1.3	149918	2	AC095571	AC095571 Rattus no
278	1.3	105604	9	AC084013	AC011345 Homo sapi	351	1.3	149953	2	AC016722	AC016722 Homo sapi
279	1.3	105797	9	AC011345	AL390316 Human DNA	352	1.3	149995	9	AL590559	AL590559 Human DNA
280	1.3	106341	9	AL390316	AL358235 Human DNA	353	1.3	150651	2	AC016230	AC016230 Homo sapi
281	1.3	106506	9	AL358235	AL121787 Human DNA	354	1.3	150747	2	AL591594	AL591594 Human DNA
282	1.3	106556	9	HS0512E2	AC011173 Homo sapi	355	1.3	152313	9	AL354830	AL354830 Human DNA
283	1.3	106676	2	AC011173	AC094530 Rattus no	356	1.3	152555	9	AC068597	AC068597 Homo sapi
284	1.3	106997	2	AC094530	AC011481 Homo sapi	357	1.3	153180	9	AC026418	AC026418 Homo sapi
285	1.3	107567	9	AC011481	AL353138 Human DNA	358	1.3	153198	2	AL365337	AL365337 Homo sapi
286	1.3	108967	9	AL353138	AC044817 Homo sapi	359	1.3	153547	2	AC004926	AC004926 Homo sapi
287	1.3	109398	9	AC044817	AL031774 Human DNA	360	1.3	153556	9	AC008567	AC008567 Homo sapi
288	1.3	109488	2	HS298J15	Continuation (2 of	361	1.3	153704	9	AC008068	AC008068 Homo sapi
289	1.3	110000	2	AC003656_1	Continuation (4 of	362	1.3	154036	2	AP002472	AP002472 Homo sapi
290	1.3	110000	2	AC073702_3	Continuation (2 of	363	1.3	154120	2	AC106852	AC106852 Homo sapi
291	1.3	110000	2	AC010536_1	Continuation (2 of	364	1.3	154452	2	AC022019	AC022019 Homo sapi
292	1.3	110000	2	AL355364_0	AL355364 Homo sapi	365	1.3	154679	2	AC022019	AC022019 Homo sapi

366	20	1.3	155124	2	AL356112	AL356112 Homo sapi	c 439	20	1.3	169159	2	AC023035	AC023035 Homo sapi
367	20	1.3	155157	9	AL360013	AL360013 Human DNA	440	20	1.3	169494	2	AC010299	AC010299 Homo sapi
368	20	1.3	155584	9	CNS07ERO	AL513143 Homo chr	441	20	1.3	169546	2	AC004157	AC004157 Plasmodiu
369	20	1.3	155666	9	AC008696	Homo sapi	442	20	1.3	169890	2	AL662872	AL662872 Homo sapi
370	20	1.3	155767	9	AP001603	Homo sapi	c 443	20	1.3	170071	2	AC094342	AC094342 Rattus no
371	20	1.3	155845	2	AC092122	Homo sapi	c 444	20	1.3	170105	3	AC007756	AC007756 Drosophill
372	20	1.3	155867	2	AC092122	Homo sapi	c 445	20	1.3	170105	3	AC009457	AC009457 Drosophill
373	20	1.3	156060	2	AC004153	AL031661 Human DNA	c 446	20	1.3	170249	3	AC006879	AC006879 Homo sapi
374	20	1.3	156499	2	HS279F22	AL62879 Homo sapi	c 447	20	1.3	170270	2	AP001548	AP001548 Homo sapi
375	20	1.3	157015	2	AP001365	AL601365 Homo sapi	c 448	20	1.3	170392	2	AC021523	AC021523 Homo sapi
376	20	1.3	157228	2	AC091485	AC091485 Homo sapi	c 449	20	1.3	170413	2	AC090405	AC090405 Homo sapi
377	20	1.3	157336	2	AC079787	AC079787 Homo sapi	c 450	20	1.3	170513	2	AC090213	AC090213 Homo sapi
378	20	1.3	157394	2	AL138790	AL138790 Homo sapi	451	20	1.3	170740	9	HSB80K6	AL121902 Human DNA
379	20	1.3	157439	2	AC026355	AC026355 Homo sapi	452	20	1.3	170760	2	AC004071	AC004071 Homo sapi
380	20	1.3	157454	2	AC037458	AC037458 Homo sapi	c 453	20	1.3	171016	2	AC018357	AC018357 Homo sapi
381	20	1.3	157485	2	AC079056	AC079056 Homo sapi	c 454	20	1.3	171069	2	AC109559	AC109559 Homo sapi
382	20	1.3	157508	9	AC018422	AC018422 Homo sapi	c 455	20	1.3	171386	2	AP001198	AP001198 Homo sapi
383	20	1.3	157711	9	AC055738	AC055738 Homo sapi	c 456	20	1.3	171790	2	AC011106	AC011106 Homo sapi
384	20	1.3	157848	9	AC079763	AC079763 Homo sapi	c 457	20	1.3	172112	2	AC060822	AC060822 Homo sapi
385	20	1.3	158143	9	AL157384	AL157384 Human DNA	c 458	20	1.3	172138	2	AC064796	AC064796 Homo sapi
386	20	1.3	158739	2	AC027402	AC027402 Homo sapi	c 459	20	1.3	172437	8	AP003301	AP003301 Oryza sat
387	20	1.3	158886	2	AC016944	AC016944 Homo sapi	c 460	20	1.3	172464	2	AC105180	AC105180 Homo sapi
388	20	1.3	159287	9	AP003684	AP003684 Homo sapi	461	20	1.3	172759	9	AC025589	AC025589 Homo sapi
389	20	1.3	159709	2	AC078806	AC078806 Homo sapi	462	20	1.3	173211	9	AL353732	AL353732 Human DNA
390	20	1.3	160114	2	AL157780	AL157780 Human DNA	463	20	1.3	173292	2	CNS01DVA	AL133502 Human chr
391	20	1.3	160391	2	AC010786	AC010786 Homo sapi	c 464	20	1.3	173913	2	AC079628	AC079628 Homo sapi
392	20	1.3	160525	2	AC027074	AC027074 Homo sapi	c 465	20	1.3	174232	2	AC026251	AC026251 Homo sapi
393	20	1.3	160556	2	AL355984	AL355984 Human DNA	c 466	20	1.3	174576	2	AC105180	AC105180 Homo sapi
394	20	1.3	160559	2	AC022807	AC022807 Homo sapi	467	20	1.3	174640	2	AC024479	AC024479 Homo sapi
395	20	1.3	161133	2	AC019285	AC019285 Homo sapi	c 468	20	1.3	174645	2	AC091111	AC091111 Homo sapi
396	20	1.3	161164	2	AC022278	AC022278 Homo sapi	c 469	20	1.3	174657	3	AC110071	AC018479 Drosophill
397	20	1.3	161298	2	AP002858	AP002858 Homo sapi	c 470	20	1.3	174902	2	AC110071	AC110071 Homo sapi
398	20	1.3	161359	2	AC090025	AC090025 Homo sapi	471	20	1.3	175218	2	AC027031	AC027031 Homo sapi
399	20	1.3	161624	2	AP001493	AP001493 Homo sapi	472	20	1.3	175353	3	AC023708	AC023708 Drosophill
400	20	1.3	161778	2	AC097003	AC097003 Rattus no	473	20	1.3	175365	2	AC069435	AC069435 Homo sapi
401	20	1.3	162064	2	AL390882	AL390882 Human DNA	474	20	1.3	175515	2	AC098128	AC098128 Rattus no
402	20	1.3	162260	2	AC021236	AC021236 Homo sapi	c 475	20	1.3	175770	9	AC026523	AC026523 Homo sapi
403	20	1.3	162539	2	AL390720	AL390720 Homo sapi	c 476	20	1.3	175850	2	AP001399	AP001399 Homo sapi
404	20	1.3	162752	9	AC099559	AC099559 Homo sapi	c 477	20	1.3	175916	2	AC007747	AC007747 Homo sapi
405	20	1.3	162837	2	AL451137	AL451137 Human DNA	478	20	1.3	176161	2	AC021173	AC021173 Homo sapi
406	20	1.3	163314	9	AL590557	AL590557 Homo sapi	c 479	20	1.3	176300	2	AC026311	AC026311 Homo sapi
407	20	1.3	163386	9	AC021659	AC021659 Homo sapi	480	20	1.3	176528	2	AC074187	AC074187 Homo sapi
408	20	1.3	163397	2	CNS01RG2	AL157690 Homo sapi	481	20	1.3	176668	2	AC027096	AC027096 Homo sapi
409	20	1.3	163537	2	AC080014	AC080014 Homo sapi	c 482	20	1.3	176734	2	AC094887	AC094887 Rattus no
410	20	1.3	163590	2	AC083781	AC083781 Homo sapi	c 483	20	1.3	176783	9	AL139281	AL139281 Human DNA
411	20	1.3	163954	2	AC090959	AC090959 Homo sapi	c 484	20	1.3	177037	9	AL163542	AL163542 Human DNA
412	20	1.3	164173	2	AC016072	AC016072 Homo sapi	c 485	20	1.3	177338	2	AC095133	AC095133 Rattus no
413	20	1.3	164452	9	CNS01RGP	AL159140 Human chr	486	20	1.3	177444	2	AC084238	AC084238 Mus muscu
414	20	1.3	164507	2	AC012067	AC012067 Homo sapi	487	20	1.3	177464	9	AP002853	AP002853 Homo sapi
415	20	1.3	165556	2	AC087190	AC087190 Homo sapi	488	20	1.3	177536	2	AC010686	AC010686 Homo sapi
416	20	1.3	165873	9	AC066593	AC066593 Homo sapi	c 489	20	1.3	177646	2	AC026317	AC026317 Homo sapi
417	20	1.3	165924	2	AC107024	AC107024 Homo sapi	c 490	20	1.3	177777	9	AC007342	AC007342 Homo sapi
418	20	1.3	166399	2	AC012584	AC012584 Homo sapi	c 491	20	1.3	177984	9	AC015700	AC015700 Homo sapi
419	20	1.3	166892	9	AC007463	AC007463 Homo sapi	c 492	20	1.3	178053	2	AC021112	AC021112 Homo sapi
420	20	1.3	166952	2	AC073856	AC073856 Homo sapi	c 493	20	1.3	178201	2	AC025888	AC025888 Homo sapi
421	20	1.3	167014	2	AC093705	AC093705 Rattus no	c 494	20	1.3	178203	9	AC008780	AC008780 Homo sapi
422	20	1.3	167478	2	AC011148	AC011148 Homo sapi	495	20	1.3	178526	2	AC106037	AC106037 Homo sapi
423	20	1.3	167512	2	AC015708	AC015708 Homo sapi	c 496	20	1.3	178637	2	AC099337	AC099337 Homo sapi
424	20	1.3	167568	2	AC063975	AC063975 Homo sapi	c 497	20	1.3	178717	2	AC024396	AC024396 Homo sapi
425	20	1.3	167589	2	AC023235	AC023235 Homo sapi	c 498	20	1.3	178756	2	AC104440	AC104440 Homo sapi
426	20	1.3	167642	9	CNS07ECT	AL442163 Human chr	499	20	1.3	179137	2	AC016641	AC016641 Homo sapi
427	20	1.3	167910	2	AC023203	AC023203 Homo sapi	c 500	20	1.3	179141	2	AC019091	AC019091 Homo sapi
428	20	1.3	167962	2	AC073484	AC073484 Homo sapi	c 501	20	1.3	179420	2	AL592296	AL592296 Homo sapi
429	20	1.3	167996	9	AC021753	AC021753 Homo sapi	c 502	20	1.3	179698	2	AC021848	AC021848 Homo sapi
430	20	1.3	168193	9	CNS05FTE2	AL359397 Human chr	c 503	20	1.3	179774	2	AC026223	AC026223 Homo sapi
431	20	1.3	168208	2	AL646091	AL646091 Mus muscu	504	20	1.3	179892	3	AC023651	AC023651 Drosophill
432	20	1.3	168265	2	AC105919	AC105919 Homo sapi	505	20	1.3	179929	2	AC006270	AC006270 Homo sapi
433	20	1.3	168306	2	AL627308	AL627308 Homo sapi	506	20	1.3	180034	2	AC072053	AC072053 Homo sapi
434	20	1.3	168571	3	AC104510	AC104510 Drosophill	507	20	1.3	180109	2	AC026896	AC026896 Homo sapi
435	20	1.3	168703	2	AC091759	Sus scrofe	c 508	20	1.3	180533	2	AL589943	AL589943 Homo sapi
436	20	1.3	168768	2	AL672047	AL672047 Mus muscu	509	20	1.3	180673	10	AC068627	AC068627 Mus Muscu
437	20	1.3	168818	2	AC013496	AC013496 Homo sapi	c 510	20	1.3	180778	9	AL355596	AL355596 Human DNA
438	20	1.3	169114	2	AC096867	AC096867 Rattus no	c 511	20	1.3	180849	30	AC023996	AC023996 Homo sapi

512	20	1.3	181148	2	AC018413	Homo sapi	585	20	1.3	199841	2	AC099729	Mus muscu	AC099729	Mus muscu
513	20	1.3	181352	9	AC010776	Homo sapi	c 586	20	1.3	199892	2	AC020997	Homo sapi	AC020997	Homo sapi
514	20	1.3	181348	2	AC023047	Homo sapi	587	20	1.3	199988	2	AC023545	Homo sapi	AC023545	Homo sapi
515	20	1.3	181563	2	AL670035	Mus muscu	588	20	1.3	200780	2	AC098827	Homo sapi	AC098827	Homo sapi
516	20	1.3	181716	2	AC068285	Homo sapi	589	20	1.3	200789	2	AL672090	Mus muscu	AL672090	Mus muscu
517	20	1.3	181823	2	AC040167	Homo sapi	590	20	1.3	201408	9	AC090644	Homo sapi	AC090644	Homo sapi
518	20	1.3	181935	2	AC079122	Homo sapi	c 591	20	1.3	201451	9	AC079623	Homo sapi	AC079623	Homo sapi
519	20	1.3	181999	2	AC106519	Rattus no	c 592	20	1.3	201579	2	AC068773	Homo sapi	AC068773	Homo sapi
520	20	1.3	182048	2	AC093293	Homo sapi	c 593	20	1.3	201757	10	AC061963	Mus muscu	AC061963	Mus muscu
521	20	1.3	182102	2	AC097274	Mus muscu	c 594	20	1.3	202423	2	AL604025	Mus muscu	AL604025	Mus muscu
522	20	1.3	182103	2	AC016998	Homo sapi	c 595	20	1.3	202550	2	AL391832	Homo sapi	AL391832	Homo sapi
523	20	1.3	182321	2	AP001357	Homo sapi	596	20	1.3	202638	2	AC109825	Homo sapi	AC109825	Homo sapi
524	20	1.3	182588	2	AC069514	Homo sapi	c 597	20	1.3	203523	2	AC068127	Homo sapi	AC068127	Homo sapi
525	20	1.3	182834	2	AC027523	Homo sapi	c 598	20	1.3	203842	9	AC010789	Homo sapi	AC010789	Homo sapi
526	20	1.3	183424	2	AC093387	Homo sapi	c 599	20	1.3	203853	2	AC098649	Homo sapi	AC098649	Homo sapi
527	20	1.3	183439	3	AC007809	Drosophi	c 600	20	1.3	203946	2	AC069465	Mus muscu	AC069465	Mus muscu
528	20	1.3	184043	9	AC097505	Homo sapi	c 601	20	1.3	204153	9	AC074091	Homo sapi	AC074091	Homo sapi
529	20	1.3	184092	2	AC018715	Homo sapi	c 602	20	1.3	204340	9	AC091103	Homo sapi	AC091103	Homo sapi
530	20	1.3	184454	3	AC104626	Drosophi	c 603	20	1.3	204767	2	AC099781	Homo sapi	AC099781	Homo sapi
531	20	1.3	184455	9	AC010740	Homo sapi	c 604	20	1.3	205815	2	AC008759	Homo sapi	AC008759	Homo sapi
532	20	1.3	184517	2	AC091758	Sus scrof	605	20	1.3	205831	2	AC074192	Homo sapi	AC074192	Homo sapi
533	20	1.3	184568	2	AC009558	Homo sapi	606	20	1.3	205932	2	AC073788	Homo sapi	AC073788	Homo sapi
534	20	1.3	184643	2	AC087650	Homo sapi	c 607	20	1.3	205993	2	AC071952	Homo sapi	AC071952	Homo sapi
535	20	1.3	184663	9	AC012113	Homo sapi	608	20	1.3	206743	2	AL663081	Mus muscu	AL663081	Mus muscu
536	20	1.3	184689	2	AL442125	Human DNA	609	20	1.3	207439	2	AC110615	Homo sapi	AC110615	Homo sapi
537	20	1.3	184916	2	AC068998	Mus muscu	c 610	20	1.3	207897	2	AL590389	Mus muscu	AL590389	Mus muscu
538	20	1.3	185024	2	AC027649	Mus muscu	611	20	1.3	207945	9	CNS01D5S	Human chr	AL121841	Human chr
539	20	1.3	185414	2	AL355526	Homo sapi	612	20	1.3	209512	9	CNS01BOV	Human chr	AL109738	Human chr
540	20	1.3	185602	2	AC097585	Sus scrof	613	20	1.3	210107	9	AL365214	Human DNA	AL365214	Human DNA
541	20	1.3	186047	2	AC074257	Homo sapi	c 614	20	1.3	211627	9	AL161651	Human DNA	AL161651	Human DNA
542	20	1.3	186150	2	AP001544	Homo sapi	c 615	20	1.3	212506	2	AC079516	Mus muscu	AC079516	Mus muscu
543	20	1.3	186156	2	AL391279	Homo sapi	c 616	20	1.3	213204	2	AC078895	Mus muscu	AC078895	Mus muscu
544	20	1.3	186656	2	AL592143	Homo sapi	c 617	20	1.3	213901	2	AC093472	Mus muscu	AC093472	Mus muscu
545	20	1.3	186716	2	AC025863	Homo sapi	618	20	1.3	214567	2	AC084316	Mus muscu	AC084316	Mus muscu
546	20	1.3	186716	3	AC023688	Drosophi	c 619	20	1.3	214696	2	AF235101	Homo sapi	AF235101	Homo sapi
547	20	1.3	187774	2	AC025586	Mus muscu	620	20	1.3	214875	10	AC015584	Mus muscu	AC015584	Mus muscu
548	20	1.3	187806	2	AC076971	Homo sapi	c 621	20	1.3	215196	2	AL603705	Mus muscu	AL603705	Mus muscu
549	20	1.3	187852	2	AC107293	Homo sapi	622	20	1.3	215653	2	AC073742	Homo sapi	AC073742	Homo sapi
550	20	1.3	188582	2	AC013720	Homo sapi	623	20	1.3	216200	2	AP000764	Mus muscu	AP000764	Mus muscu
551	20	1.3	188863	9	AC007345	Homo sapi	624	20	1.3	216766	2	AC103953	Homo sapi	AC103953	Homo sapi
552	20	1.3	189134	9	H0K0004682	Homo sapi	625	20	1.3	216988	10	AC084217	Mus Muscu	AC084217	Mus Muscu
553	20	1.3	189427	2	AC023286	Mus muscu	626	20	1.3	217421	9	AC006211	Homo sapi	AC006211	Homo sapi
554	20	1.3	189801	30	AC027447	Homo sapi	c 627	20	1.3	217730	2	AC087440	Homo sapi	AC087440	Homo sapi
555	20	1.3	189893	2	AP000999	Homo sapi	c 628	20	1.3	217747	2	AC105448	Homo sapi	AC105448	Homo sapi
556	20	1.3	189903	2	AC012542	Homo sapi	629	20	1.3	218589	2	AC025083	Homo sapi	AC025083	Homo sapi
557	20	1.3	190856	2	AC009160	Homo sapi	630	20	1.3	219175	9	AC092641	Homo sapi	AC092641	Homo sapi
558	20	1.3	190856	2	AC024505	Homo sapi	631	20	1.3	219720	2	AC023364	Mus muscu	AC023364	Mus muscu
559	20	1.3	191464	9	CNS01DPI	Human chr	c 632	20	1.3	221409	3	AE003730	Drosophi	AE003730	Drosophi
560	20	1.3	191485	2	AC091467	Mus muscu	633	20	1.3	221688	2	AE010323	Homo sapi	AE010323	Homo sapi
561	20	1.3	191655	9	AC020661	Homo sapi	c 634	20	1.3	221830	3	AC009393	Homo sapi	AC009393	Homo sapi
562	20	1.3	191911	9	AC098965	Homo sapi	c 635	20	1.3	222329	10	AC074310	Mus muscu	AC074310	Mus muscu
563	20	1.3	191911	9	AC098965	Homo sapi	636	20	1.3	224010	2	AP001846	Homo sapi	AP001846	Homo sapi
564	20	1.3	192179	2	AC097365	Mus muscu	637	20	1.3	226199	2	AC084069	Mus muscu	AC084069	Mus muscu
565	20	1.3	192189	2	AC107212	Homo sapi	c 638	20	1.3	226323	3	AE003697	Drosophi	AE003697	Drosophi
566	20	1.3	192358	3	AC099028	Homo sapi	639	20	1.3	226825	2	AC073912	Homo sapi	AC073912	Homo sapi
567	20	1.3	192581	2	PFMAL13P1	Plasmodiu	c 640	20	1.3	227348	2	AC092851	Homo sapi	AC092851	Homo sapi
568	20	1.3	192819	2	AC012280	Homo sapi	c 641	20	1.3	228156	9	AC022150	Homo sapi	AC022150	Homo sapi
569	20	1.3	193098	9	AL137140	Human DNA	c 642	20	1.3	229655	3	AE003370	Drosophi	AE003370	Drosophi
570	20	1.3	193117	2	AC106721	Homo sapi	c 643	20	1.3	231661	2	AC073770	Mus muscu	AC073770	Mus muscu
571	20	1.3	194041	2	AL662912	Mus muscu	644	20	1.3	232744	3	AE003705	Drosophi	AE003705	Drosophi
572	20	1.3	194198	2	AC011774	Homo sapi	645	20	1.3	234235	2	AC099303	Rattus no	AC099303	Rattus no
573	20	1.3	194418	9	AF235098	Homo sapi	646	20	1.3	235150	2	AC073776	Mus muscu	AC073776	Mus muscu
574	20	1.3	194487	2	AP002896	Homo sapi	c 647	20	1.3	237119	2	AC017740	Drosophi	AC017740	Drosophi
575	20	1.3	195102	9	AL353795	Human DNA	c 648	20	1.3	239108	2	AL606661	Mus muscu	AL606661	Mus muscu
576	20	1.3	195250	2	AC023607	Mus muscu	649	20	1.3	240294	2	AC044892	Homo sapi	AC044892	Homo sapi
577	20	1.3	195250	2	AC023607	Mus muscu	c 650	20	1.3	241857	2	AL645911	Mus muscu	AL645911	Mus muscu
578	20	1.3	195743	2	AC016774	Homo sapi	651	20	1.3	242965	2	AC023175	Mus muscu	AC023175	Mus muscu
579	20	1.3	196349	2	AC019112	Homo sapi	652	20	1.3	247337	2	AC099070	Rattus no	AC099070	Rattus no
580	20	1.3	196413	9	AC097269	pen trogl	c 653	20	1.3	258174	2	AC079429	Mus muscu	AC079429	Mus muscu
581	20	1.3	198628	2	AC007363	Homo sapi	654	20	1.3	259978	2	AC073819	Mus muscu	AC073819	Mus muscu
582	20	1.3	198829	2	AL391668	Homo sapi	655	20	1.3	261090	3	AE003798	Drosophi	AE003798	Drosophi
583	20	1.3	199377	2	AC084294	Mus muscu	656	20	1.3	265815	2	AC027308	Homo sapi	AC027308	Homo sapi
584	20	1.3	199730	2	AC083940	Homo sapi	657	20	1.3	279090	2	AC091290	Mus muscu	AC091290	Mus muscu

c 658	20	1.3	286060	2	AL589862	Homo sapi	731	19	1.2	630	8	PRU90342	U90342 Pinus radia
659	20	1.3	299537	3	AE003438	Drosophill	732	19	1.2	631	3	HVY300658	AJ00658 Heliothis
660	20	1.3	301051	3	AE003487	Drosophill	733	19	1.2	637	9	HUWY51E04	AJ08804 Homo sapi
c 661	20	1.3	302350	3	AE003485	Drosophill	734	19	1.2	647	3	AF420271	AF420271 Culex pip
662	20	1.3	340000	9	AP001698	Homo sapi	735	19	1.2	653	3	PU33869	U33869 Plasmodium
c 663	20	1.3	340000	9	AP001710	Homo sapi	736	19	1.2	662	3	HS438189	AJ338189 Homo sapi
664	20	1.3	340000	9	AP001731	Homo sapi	c 737	19	1.2	663	11	G383327	G38327 RPTC-6-190H
c 665	20	1.3	340000	9	AP001753	Homo sapi	c 738	19	1.2	673	3	HS435190	HS435190 Homo sapi
c 666	20	1.3	349980	6	AX344555	Sequence	c 739	19	1.2	674	8	HS435190	AJ35190 Drosophill
c 667	20	1.3	349980	6	AX344560	Sequence	c 740	19	1.2	678	8	AY052654	AY052654 Arabidops
c 668	20	1.3	349980	6	AX344561	Sequence	c 741	19	1.2	678	9	BC004274	BC004274 Homo sapi
c 669	20	1.3	349980	6	AX344572	Sequence	742	19	1.2	683	9	AT143870	AT143870 Homo sapi
c 670	19	1.2	126	6	MTDUCPULA	Sequence	743	19	1.2	687	3	AY071700	AY071700 Drosophill
c 671	19	1.2	147	11	G26109	Sequence	c 744	19	1.2	687	9	HS434845	AJ34845 Homo sapi
c 672	19	1.2	148	6	AX187356	Sequence	c 745	19	1.2	695	10	RATCARB06	M23952 Rat carboxy
c 673	19	1.2	149	6	DME426863	Sequence	746	19	1.2	719	9	BC007857	BC007857 Homo sapi
674	19	1.2	162	8	PERCPRABX	Sequence	747	19	1.2	728	10	BC021589	BC021589 Mus muscu
675	19	1.2	165	8	NBCPULA	Sequence	748	19	1.2	732	8	AF031543	AF031543 Fritilliar
676	19	1.2	182	3	AF010512	Sequence	749	19	1.2	734	8	PAB131105	AJ131105 Picea abi
677	19	1.2	191	8	NGCPULA	Sequence	750	19	1.2	744	9	BC007410	BC007410 Homo sapi
678	19	1.2	191	8	NGCPULA	Sequence	751	19	1.2	744	9	BC007434	BC007434 Homo sapi
679	19	1.2	193	8	NVECPULA	Sequence	c 752	19	1.2	748	8	AF207566	AF207566 Brassica
680	19	1.2	195	8	NGICPULA	Sequence	c 753	19	1.2	756	8	AF207572	AF207572 Brassica
681	19	1.2	203	8	NMCPULA	Sequence	754	19	1.2	756	8	AF322226	AF322226 Lupinus l
682	19	1.2	205	8	NATPULA	Sequence	755	19	1.2	766	8	LEACIDIN	X77264 L.esculentu
683	19	1.2	254	6	AX187664	Sequence	756	19	1.2	793	3	AF117600	AF117600 Manduca s
684	19	1.2	262	3	AF134815	Sequence	757	19	1.2	794	9	BC008643	BC008643 Homo sapi
685	19	1.2	262	11	G62158	Sequence	758	19	1.2	803	8	AF244687	AF244687 Zea mays
686	19	1.2	292	11	G30913	Sequence	759	19	1.2	807	3	AY071462	AY071462 Drosophill
687	19	1.2	336	11	AF346024	Sequence	760	19	1.2	811	8	AB028153	AB028153 Prunus av
c 688	19	1.2	338	6	AX106357	Sequence	761	19	1.2	821	6	AF284786	AF284786 Aedes alb
c 689	19	1.2	338	6	AX140648	Sequence	762	19	1.2	821	6	AX187075	AX187075 Sequence
c 690	19	1.2	338	6	AX200508	Sequence	763	19	1.2	822	8	ALFEESIPC	M7189 Medicago sa
c 691	19	1.2	338	6	AX267164	Sequence	764	19	1.2	823	9	BC020914	BC020914 Homo sapi
c 692	19	1.2	338	11	G23737	Sequence	765	19	1.2	825	8	AY064032	AY064032 Arabidops
693	19	1.2	345	6	AX260663	Sequence	766	19	1.2	829	8	SA0947	SA0947 abscisic ac
c 694	19	1.2	363	6	AX070539	Sequence	767	19	1.2	837	9	HS4326732	AJ226732 Homo sapi
c 695	19	1.2	374	6	AX340671	Sequence	768	19	1.2	838	10	BC008229	BC008229 Mus muscu
c 696	19	1.2	375	3	SCU71450	Sequence	769	19	1.2	839	8	AY050903	AY050903 Arabidops
697	19	1.2	385	3	AF288608	Sequence	c 770	19	1.2	842	3	DDU49170	U49170 Dictyostell
c 698	19	1.2	393	6	AX071209	Sequence	771	19	1.2	854	5	AB028455	AB028455 Cyprinus
699	19	1.2	408	3	AF288607	Sequence	772	19	1.2	855	8	CAR275317	AJ275317 Cicer ari
c 700	19	1.2	412	11	DM161B3T	Sequence	773	19	1.2	866	8	PS4316577	AJ316577 Pisum sat
701	19	1.2	421	8	MTSCPULB	Sequence	774	19	1.2	867	9	BC015482	BC015482 Homo sapi
702	19	1.2	433	6	NBCPULB	Sequence	775	19	1.2	869	9	AK000193	AK000193 Homo sapi
703	19	1.2	437	6	AX186591	Sequence	776	19	1.2	885	8	ATKZAP10	ATKZAP10 A.thaliana
c 704	19	1.2	437	6	AX331782	Sequence	777	19	1.2	887	3	AY058299	AY058299 Drosophill
c 705	19	1.2	437	6	AX337665	Sequence	c 778	19	1.2	905	10	MMU277212	AJ277212 Mus muscu
706	19	1.2	460	11	G24674	Sequence	779	19	1.2	908	3	AY075321	AY075321 Drosophill
c 707	19	1.2	466	5	XLSUMO	Sequence	780	19	1.2	915	9	IR0200247	AL109727 Homo sapi
708	19	1.2	472	3	AF334182	Sequence	781	19	1.2	922	8	HR8223388	AJ223388 Heyera bra
709	19	1.2	499	8	AY065450	Sequence	782	19	1.2	923	3	AY071632	AY071632 Drosophill
710	19	1.2	500	3	TML27APR	Sequence	783	19	1.2	931	6	AR123131	AR123131 Sequence
c 711	19	1.2	501	6	AX255932	Sequence	784	19	1.2	931	6	AR124083	AR124083 Sequence
712	19	1.2	518	11	G30584	Sequence	785	19	1.2	931	6	AF243361	AF243361 Glycine m
713	19	1.2	528	8	HDMR301R	Sequence	786	19	1.2	933	8	AF017991	AF017991 Arabidops
c 714	19	1.2	530	6	AA3742	Sequence	787	19	1.2	933	6	ATU75601	ATU75601 Arabidops
715	19	1.2	530	6	AR071495	Sequence	c 788	19	1.2	936	6	AX186156	AX186156 Sequence
c 716	19	1.2	530	11	G55823	Sequence	789	19	1.2	941	9	FSY420188	FSY420188 Fagus syl
c 717	19	1.2	539	6	AX184493	Sequence	790	19	1.2	941	9	HSM800591	AJ080090 Homo sapi
718	19	1.2	540	3	AF400188	Sequence	791	19	1.2	951	3	AY069243	AY069243 Drosophill
719	19	1.2	543	6	AX188166	Sequence	792	19	1.2	959	8	AF230710	AF230710 Chloranth
c 720	19	1.2	555	6	AX237655	Sequence	793	19	1.2	965	3	AY071499	AY071499 Drosophill
721	19	1.2	558	8	TAECM3	Sequence	794	19	1.2	980	3	BC014603	BC014603 Homo sapi
722	19	1.2	577	8	ATU75204	Sequence	795	19	1.2	983	3	TYU87098	TYU87098 Trichomonas
723	19	1.2	599	8	TAEPM1	Sequence	796	19	1.2	984	4	AF059493	AF059493 Sus scrof
724	19	1.2	601	8	AT045861	Sequence	797	19	1.2	985	8	AB028229	AB028229 Arabidops
c 725	19	1.2	603	8	BC001204	Sequence	798	19	1.2	992	8	DIA297612	AJ297612 Digitalis
726	19	1.2	610	8	AF239929	Sequence	799	19	1.2	994	8	AY045920	AY045920 Arabidops
727	19	1.2	613	10	BC002097	Sequence	800	19	1.2	1001	6	AS9875	AS9875 Sequence 4
728	19	1.2	615	9	HS4248265	Sequence	801	19	1.2	1004	5	DRAPIRPAI	AF12653 Danio rerio
729	19	1.2	615	9	HS4248265	Sequence	802	19	1.2	1006	8	AF126737	AF126737 Phaseolus
730	19	1.2	616	3	DDALPHA1	Sequence	803	19	1.2	1006	8	AF129686	AF129686 Vigna rad

804	19	1.2	1008	8	BNPBP4	X90730 B. napsus mRN	877	19	1.2	1403	8	ATU83624	U83624 Arabidopsis
805	19	1.2	1009	3	AF030628	AF030628 Plasmodiu	878	19	1.2	1414	8	ATH7587	AF007587 Arabidops
806	19	1.2	1011	3	AY075365	AY075365 Drosophi	879	19	1.2	1421	10	AF051425	AF051425 Rattus no
807	19	1.2	1021	9	BC005200	BC005200 Homo sapi	880	19	1.2	1439	8	AY065232	AY065232 Arabidops
808	19	1.2	1035	9	BC002828	BC002828 Homo sapi	881	19	1.2	1453	3	AY058384	AY058384 Drosophi
809	19	1.2	1036	8	BVCHLORP	Y13865 Beta vulgar	882	19	1.2	1457	9	BC003017	BC003017 Homo sapi
810	19	1.2	1041	3	AF231352	AF231352 Ctenoceph	883	19	1.2	1458	9	AF070645	AF070645 Homo sapi
811	19	1.2	1045	8	AY034967	AY034967 Arabidops	884	19	1.2	1466	9	BC000460	BC000460 Homo sapi
812	19	1.2	1047	8	BC005132	BC005132 Homo sapi	885	19	1.2	1477	8	AF028842	AF028842 Arabidops
813	19	1.2	1048	8	AF297521	AF297521 Prunus av	886	19	1.2	1481	3	AY058287	AY058287 Drosophi
814	19	1.2	1054	10	BC021770	BC021770 Mus muscu	887	19	1.2	1482	5	AB071891	AB071891 Letheuter
815	19	1.2	1065	3	AY061524	AY061524 Drosophi	888	19	1.2	1507	9	AF090988	AF090988 Homo sapi
816	19	1.2	1080	8	NTT232AF	X62427 N. tabacum m	889	19	1.2	1512	8	AB064519	AB064519 Oryza sat
817	19	1.2	1085	5	AF270790	AF270790 Orosachrom	890	19	1.2	1516	9	BC020631	BC020631 Homo sapi
818	19	1.2	1093	8	AB047397	AB047397 Zygosaech	891	19	1.2	1531	9	BC001494	BC001494 Homo sapi
819	19	1.2	1096	9	BC004277	BC004277 Homo sapi	892	19	1.2	1541	8	AF370323	AF370323 Arabidops
820	19	1.2	1102	9	BC012033	BC012033 Homo sapi	893	19	1.2	1542	6	AX001335	AX001335 Sequence
821	19	1.2	1107	9	HSMB02425	AL157420 Homo sapi	894	19	1.2	1546	5	AF102890	AF102890 Gadus mor
822	19	1.2	1110	5	SSAJ4825	AF004825 Salmo sal	895	19	1.2	1556	9	BC008009	BC008009 Homo sapi
823	19	1.2	1113	8	AF043519	AF043519 Arabidops	896	19	1.2	1566	9	BC004983	BC004983 Homo sapi
824	19	1.2	1114	9	BC015039	BC015039 Homo sapi	897	19	1.2	1569	9	AK026099	AK026099 Homo sapi
825	19	1.2	1115	3	AY060811	AY060811 Drosophi	898	19	1.2	1570	3	AY069282	AY069282 Drosophi
826	19	1.2	1137	10	BC002003	BC002003 Mus muscu	899	19	1.2	1578	6	AK138898	AK138898 Sequence
827	19	1.2	1140	3	AY069465	AY069465 Drosophi	900	19	1.2	1578	6	E29750	E29750 cDNA clone
828	19	1.2	1150	5	AF140281	AF140281 Ciriethinus	901	19	1.2	1582	3	AY052107	AY052107 Drosophi
829	19	1.2	1160	10	AF408432	AF408432 Mus muscu	902	19	1.2	1584	9	BC002601	BC002601 Homo sapi
830	19	1.2	1161	10	BC002332	BC002332 Mus muscu	903	19	1.2	1584	10	BC016573	BC016573 Mus muscu
831	19	1.2	1163	5	SSAJ4824	AI004824 Salmo sal	904	19	1.2	1598	5	XLXNWT8	XLXNWT8 X. laevis xw
832	19	1.2	1168	3	AF165522	AF165522 Homo sapi	905	19	1.2	1631	9	BC012134	BC012134 Homo sapi
833	19	1.2	1171	3	AY075439	AY075439 Drosophi	906	19	1.2	1642	8	AY062178	AY062178 Oryza sat
834	19	1.2	1172	8	AY072075	AY072075 Arabidops	907	19	1.2	1644	8	AY072108	AY072108 Arabidops
835	19	1.2	1177	8	AY063837	AY063837 Arabidops	908	19	1.2	1648	8	AY050893	AY050893 Arabidops
836	19	1.2	1182	8	AF451952	AF451952 Arabidops	909	19	1.2	1654	8	AY056117	AY056117 Arabidops
837	19	1.2	1186	8	BC000793	BC000793 Homo sapi	910	19	1.2	1656	5	AB010713	AB010713 Cypripus
838	19	1.2	1196	4	CAU35049	U33049 Ovis aries	911	19	1.2	1659	9	BC015312	BC015312 Homo sapi
839	19	1.2	1201	9	BC012589	BC012589 Homo sapi	912	19	1.2	1665	9	AK024914	AK024914 Homo sapi
840	19	1.2	1202	9	HUMETFC4	LI18960 Human prote	913	19	1.2	1667	14	AB021977	AB021977 Rinderpes
841	19	1.2	1203	8	AY063792	AY063792 Arabidops	914	19	1.2	1684	8	AF370136	AF370136 Arabidops
842	19	1.2	1204	10	BC005743	BC005743 Mus muscu	915	19	1.2	1687	3	AY052117	AY052117 Drosophi
843	19	1.2	1205	9	BC005840	BC005840 Homo sapi	916	19	1.2	1692	9	BC010359	BC010359 Homo sapi
844	19	1.2	1215	3	AY071536	AY071536 Drosophi	917	19	1.2	1700	8	AF166352	AF166352 Arabidops
845	19	1.2	1215	8	LPE243504	AJ243504 Lolium pe	918	19	1.2	1701	9	HS421991	HS421991 Homo sapi
846	19	1.2	1223	3	DDIDISIA	J01282 Slime mold	919	19	1.2	1703	9	HSMRINTX	X44709 H. sapiens m
847	19	1.2	1223	8	GGN132215	AJ132215 Gnetum gn	920	19	1.2	1703	9	AB048936	AB048936 Macaca fa
848	19	1.2	1224	8	AY050929	AY050929 Arabidops	921	19	1.2	1704	9	H009384S20	AF219908 Homo sapi
849	19	1.2	1226	8	AK026455	AK026455 Homo sapi	922	19	1.2	1708	9	AF070559	AF070559 Homo sapi
850	19	1.2	1239	10	MMCENIRA	X17373 M. musculus	923	19	1.2	1717	10	BC023168	BC023168 Mus muscu
851	19	1.2	1242	10	AF143369	AF143369 Mus muscu	924	19	1.2	1724	9	BC001179	BC001179 Homo sapi
852	19	1.2	1245	9	AF282025	AF282025 Homo sapi	925	19	1.2	1733	9	HSOCRT6R	Z18284 Homo sapien
853	19	1.2	1249	9	BC021697	BC021697 Homo sapi	926	19	1.2	1735	9	BC017836	BC017836 Homo sapi
854	19	1.2	1250	8	AY048692	AY048692 Arabidops	927	19	1.2	1739	9	AB048980	AB048980 Homo sapi
855	19	1.2	1261	9	BC021105	BC021105 Homo sapi	928	19	1.2	1742	10	BC016107	BC016107 Mus muscu
856	19	1.2	1266	9	HSMB02163	AL137452 Homo sapi	929	19	1.2	1744	3	AK024845	AK024845 Homo sapi
857	19	1.2	1270	9	BC019326	BC019326 Homo sapi	930	19	1.2	1753	9	AY061291	AY061291 Drosophi
858	19	1.2	1273	9	BC021667	BC021667 Homo sapi	931	19	1.2	1755	9	HSAF000974	AF000974 Human zyx
859	19	1.2	1286	10	AF061262	AF061262 Mus muscu	932	19	1.2	1777	9	HSTECTXT11	U43377 Human tyros
860	19	1.2	1300	9	AF005214	AF005214 Homo sapi	933	19	1.2	1806	10	MMU295253	AJ295253 Mus muscu
861	19	1.2	1302	9	BC008906	BC008906 Homo sapi	934	19	1.2	1816	3	AF273770	AF273770 Caenorhab
862	19	1.2	1318	6	AX008671	AX008671 Sequence	935	19	1.2	1817	5	XLA011652	AX011652 Xenopus l
863	19	1.2	1318	8	FCO132438	AJ132438 Phasaelus	936	19	1.2	1819	9	AY007790	AY007790 Homo sapi
864	19	1.2	1329	8	AY045905	AY045905 Arabidops	937	19	1.2	1827	8	CRO7332	AJ007332 Catharant
865	19	1.2	1331	10	BC013440	BC013440 Mus muscu	938	19	1.2	1839	9	AK000096	AK000096 Homo sapi
866	19	1.2	1332	9	HSAA20565	AJ20565 Homo sapi	939	19	1.2	1841	3	AY060776	AY060776 Drosophi
867	19	1.2	1341	3	DME278068	AJ278068 Drosophi	940	19	1.2	1843	8	NTA271750	AJ271750 Nicotiana
868	19	1.2	1347	3	BC002596	BC002596 Homo sapi	941	19	1.2	1844	9	BC010741	BC010741 Homo sapi
869	19	1.2	1350	3	AY071533	AY071533 Drosophi	942	19	1.2	1858	9	AB046649	AB046649 Macaca fa
870	19	1.2	1362	8	AY063851	AY063851 Arabidops	943	19	1.2	1860	6	AF015566	AF015566 Dictyoste
871	19	1.2	1364	8	AK000039	AK000039 Homo sapi	944	19	1.2	1874	3	AY071374	AY071374 Drosophi
872	19	1.2	1367	8	AY065422	AY065422 Arabidops	945	19	1.2	1874	3	AY061829	AY061829 Drosophi
873	19	1.2	1371	8	AY035016	AY035016 Arabidops	946	19	1.2	1879	3	AY058766	AY058766 Drosophi
874	19	1.2	1385	3	AY058467	AY058467 Drosophi	947	19	1.2	1883	5	AB025103	AB025103 Gallus ga
875	19	1.2	1386	3	HSU77327	U77327 Human K1-1/	948	19	1.2	1889	5	AB025103	AB025103 Gallus ga
876	19	1.2	1401	9	BC000932	BC000932 Homo sapi	949	19	1.2	1889	10	BC007468	BC007468 Mus muscu

Qy 721 gccacatttcagtcagagactaactacggtacatgagatcactcagcgtcagcgg 780
 Db 721 gccacatttcagtcagagactaactacggtacatgagatcactcagcgtcagcgg 780
 Qy 781 cgtctctcatgtacacggtcgtcactacacccggtcgtcgtcgtcgtcgtcgtc 840
 Db 781 cgtctctcatgtacacggtcgtcactacacccggtcgtcgtcgtcgtcgtcgtc 840
 Qy 841 ctctcagccttcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtc 900
 Db 841 ctctcagccttcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtc 900
 Qy 901 ctatcatcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtc 960
 Db 901 ctatcatcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtc 960
 Qy 961 aacatactcactcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtc 1020
 Db 961 aacatactcactcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtc 1020
 Qy 1021 aacatactcactcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtc 1080
 Db 1021 aacatactcactcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtc 1080
 Qy 1081 gagcgcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtc 1140
 Db 1081 gagcgcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtc 1140
 Qy 1141 agcacacgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtc 1200
 Db 1141 agcacacgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtc 1200
 Qy 1201 gagcgcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtc 1260
 Db 1201 gagcgcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtc 1260
 Qy 1261 aagcgcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtc 1320
 Db 1261 aagcgcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtc 1320
 Qy 1321 tccctcagcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtc 1380
 Db 1321 tccctcagcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtc 1380
 Qy 1381 ctcaagcgcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtc 1440
 Db 1381 ctcaagcgcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtc 1440
 Qy 1441 gtaattatgt 1500
 Db 1441 gtaattatgt 1500
 Qy 1501 gacctcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtc 1539
 Db 1501 gacctcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtc 1539

RESULT 2
 AY005148 1421 bp mRNA linear INV 01-AUG-2001
 LOCUS AY005148
 DEFINITION Drosophila melanogaster nicotinic acetylcholine receptor beta
 subunit mRNA, complete cds.
 ACCESSION AY005148
 VERSION AY005148.1 GI:15072330
 KEYWORDS
 SOURCE
 ORGANISM
 Drosophila melanogaster
 Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
 Pterygota; Neoptera; Endopterygota; Diptera; Brachyera;
 Muscomorpha; Ephydroidea; Drosophilidae; Drosophila.
 REFERENCE
 1 (bases 1 to 1421)
 Grauso M, and Sattelle D.B.
 Drosophila melanogaster nicotinic acetylcholine receptor beta3

JOURNAL subunit (nacrbeta-21C gene)
 REFERENCE 2 (bases 1 to 1421)
 AUTHORS Grauso M, and Sattelle D.B.
 TITLE Direct Submission
 JOURNAL Submitted (18-JUL-2000) MRC-FGU Human Anatomy and Genetics,
 University of Oxford, South Parks Road, Oxford OX1 3QX, UK
 FEATURES
 source
 1.1421
 /organism="Drosophila melanogaster"
 /db_xref="taxon:7227"
 /chromosome="2"
 /map="2L; 21C5-21C5"
 /dev_stage="embryo"
 45.1370
 /note="nacrbeta-21C; ion channel; transmembrane
 neurolinsmiller"
 /codon_start=1
 /product="nicotinic acetylcholine receptor beta subunit"
 /protein_id="AAF89090.1"
 /db_xref="GI:15072331"
 /translation="MTTTRKAPVSGPGLPLLQMLMGLTSTVPGATATADPK
 NAWKALDRLHAGLFTNYDSVQVPEQPTNVLDMVYVYIDIDELNKLTHWLN
 LKRWDERVWOPSOYDNITQITLKSSEVTPQITLNGDEGLMAETQVTLSDGHR
 KMPPAYTVAYCELMLNMPHDKOSKLTIGSMGLKVLPENGTARDESIDHDLVOSP
 EMEIVDSRAHFVSODYVYMEYVLTQARSSSVYTAIVTASCTIVTALSAPFLPBM
 GGRKININGLIIVIAFLMFPQLLPVLSNTPVIVITSTSLTSLVSTVYEVLY
 YLATGKHKRLPEALRKLHLGLTLLSVSTSTGESQAEKTKENDEHYERADROE
 SSPLGINHVEVPAKANQDFDWALATVAVRISVSPSLAFLIALINCSV"
 BASE COUNT 311 a 463 c 389 g 258 t
 ORIGIN

Query Match 79.3%; Score 1220; DB 3; Length 1421;
 Best Local Similarity 99.8%; Pred. No. 0;
 Matches 1370; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy 14 gtatcatcgaagaaagcggtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtc 73
 Db 16 GTATCATCGAAGAAAGCGGCGGTGTAACATGACGAGACTCCCAAGATTAAGCACACG 75
 Qy 74 ttccgcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtc 133
 Db 76 TTTCGGTCTCGAGCTCCACTACTGCTGCAAAATGTAATGGGAGATGCTTATGGGGC 135
 Qy 134 tgcattcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtc 193
 Db 136 TGACTTCGGTCCAGGCGGCGCACTGCGACCGGCGACCCCAAGAACGCAATGTCAAGGCC 195
 Qy 194 tgcattcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtc 253
 Db 196 TGATTCGGTCTCCAGCGGCGGTGTTCAAGACTGACAGACAGGATGTGACGCGGTTC 255
 Qy 254 aaggaacccccaagaaagcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtc 313
 Db 256 AAGGAACCCCAAGAAAGTGTGGAATGTGAGTGTGACCTATGATGACATGACAGT 315
 Qy 314 tgaagcgaagcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtc 373
 Db 316 TGAAGCGCAAGCTGACACCCACTGCTGATCTCCGATGAGAGAGAGAGCGCG 375
 Qy 374 tctggaacgcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtc 433
 Db 376 TGTGGCAACGCTCACATATGACACATCAGCGATCATCTTGAATGCCAGGAGTCT 435
 Qy 434 ggaacccccaatcacactcttcaacgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgc 493
 Db 436 GGACCCCAATCACACTCTTCAAGCGCGCGAGAGGTGCGCTGATGCGCGAAACCGAGG 495
 Qy 494 tgcattcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtc 553
 Db 496 TGACCTTCAGCGACGATGCGACACTTCCGGTGAGATGCTCCAGCGGTGTACACGCGCTACT 555

QY	554	ggaactcaacatgctcaacttggcccaacgaacagcagctcgaatttgaagatcggc	613
Db	556	GGGAATCAACATGCTCAACTTGCCCCCAAGACAGACAGCTGCAAGTTGAAGATCGGCT	615
QY	614	ccttgaggcctgaagtgctcctccgggaacgycagcgcgagagagatcccttgacc	673
Db	616	CTTGCGGCGCTGAGAGTCCCTCCGCGGGAAGACGACGCGCGAGAGAGTCCCTTGACC	675
QY	674	acgaagaccctgttcaagcaccggatgggaatctgtgacccggagccacttgtca	733
Db	676	ACGAGACCTGGTTAGTACCGGAGTGGGAATGTGTGACTCGGAGGCCACTTGTGA	735
QY	734	gtcaagagctactacagcgtctacatgtgagtaacactctgacggtctaaagcgctccctcaagt	793
Db	736	GTCAGAGCTACTACGGCTACATGAGGTACACTCTGACGGCTCAGCGGCGCTCTCCATGT	795
QY	794	acaagcgccgtatctcaaacaccccgctcctgtatcgtatcctctgaccttcaagccttc	853
Db	796	ACACGCGCGTCACTACACACCCGGGCTCTGATGTATCTTGAGCCCTTGAGCCTTCT	855
QY	854	ggctcccccacacatggcgcgagaaatcatgatcaagcgctctcatcatcgtga	913
Db	856	GGCTGCTCTCCACATGGGGCGGCGAGATGATATACAGCGCTGCTCATATCGTGA	915
QY	914	tcgcgcctctcctaatgtactcgcacagctctctgcagctgtctgtccaactaccac	973
Db	916	TGCGCGCTTCTCTCATGTACTTCGCGGACCTCTCCAGTGTGTCCAACTACTCTCAC	975
QY	974	ttgtgtatctcttcaagcaacacgcctgtctatctgaagcttccacatcgtccgaag	1033
Db	976	TTGTGTATCTCTTACACAGCACACAGCTGCTGTATCTGAGCTCTCCACATGCTGAGG	1035
QY	1034	ttctagctctgtacccctgycacacagcacaagaagcgccctgcgcgagagcgctcgaaga	1093
Db	1036	TTCTAGTCTGTACTGTGCGCACAGGCAAGACAAAGAGGCGCTGCGGAGCGCTGAGAA	1095
QY	1094	agctgtctcaagggagacttggcgaagtgtgtctgtctcgtgtgttcagacacactggcg	1153
Db	1096	AGCTTGTACAGGGGACCTGGGACGTGGCTGTCTCTCGGTGTTCACACACCGGGG	1155
QY	1154	agtcgacggcgagagagaccaaagatctggaagcaccgcctacagagagcgagcagac	1213
Db	1156	AGTGCAGGCGGAGAGAACAAAGAGATGAGAGACACCCGTACGAGAGGCGGAGAGAC	1215
QY	1214	aggagctcagtcgcgtggtgcatcaacacacccgaagtgcggcgccaaagccaaacagt	1273
Db	1216	AGGAATCCAGTCCGCTGGGCAATCAACACACGAGAGTGGCGGCGCCAAAGGCAACACT	1275
QY	1274	tcgaactggcgctgtgtggcgaacgcgcgtggagacgcatctctctgttccctcagccttg	1333
Db	1276	TGGAATGGGCGCTGTGGGCAACCGCGTGTGAGACCGCATTTCTTCTGTTCTTCAAGCTTGG	1335
QY	1334	cccttcctatcttgccatcaagtgctcogtltagagatgtctgagactcgaactcaag	1386
Db	1336	CTTCTCATCTTGCGCATCAGGTGCTCGTGTAGGATGTCTCGAGACTCAAG	1388
RESULT 3			
DME318761		1334 bp	linear
LOCUS			
DEFINITION		Drosophila melanogaster mRNA for nicotinic acetylcholine receptor	
ACCESSION		beta 3 (Dbeta3) subunit (nacrbeta-21C gene).	
VERSION		AJ318761	
KEYWORDS		AJ318761.1 GI:15130913	
SOURCE		nacrbeta-21C gene; nicotinic acetylcholine receptor beta 3.	
ORGANISM		Fruit fly.	
REFERENCE		Drosophila melanogaster	
AUTHORS		Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;	
JOURNAL		Pterygota; Neoptera; Endopterygota; Diptera; Brachycera;	
		Muscomorpha; Ephydroidea; Drosophilidae; Drosophila.	
		1 (bases 1 to 1334)	
		Lansdel, S.J. and Millar, N.S.	
		Unpublished	


```

|||||
Db 541 CCACGACAAACAGAGCTGCAAGTTGAAGATCGGCTCTGGGGCTGAAGGTCGCTGCC 600
Qy 639 ggaagaacgagcagcgagagagagatcccttgacacgaagacgtgttcaatcacgga 638
Db 601 GGAGAACGGACGGCGGAGAGAGATCCCTTGACCAAGACGACCTGGTTAGTCACCGGA 660
Qy 699 gtgggaatcgttgactcgcgcgagccactgttcagtcagactactaagcgtacatgga 758
Db 661 GTGGGAATATCTGGACTCGGAGCCCACTTTGTACGTACGAGACTACTACGGCTACATGA 720
Qy 759 gtacacttgcagcgcgtcgcctccatcatgtaacagcgccgtatcatcacaccgc 818
Db 721 GTACACTGTACGGCTCAGCGGGCTCCTTCATGATACAGGGCGGTATCTACACACCGC 780
Qy 819 gtccatcatcgtcatccctgcgccttcagccttcagccttcagccttcagccttcag 878
Db 781 GTCTGCTATCTGATCTGCGCTCTCAGCCTTCTGCTCTGCTCTGCTCTGCTCTGCTCTG 840
Qy 879 gaagatcatgatacaacgagcgtctcatcatcgtatcgcgccttcctcatgtaattcgc 938
Db 841 GAAGATCATATATCAACGGCTGCTCATCATGATGCGCCTTCTCTCATGTAATTGCG 900
Qy 939 ccaagctcctgcagctgtctgcacaaatactcaactgtgtgtaattcttctacagcag 998
Db 901 CCAAGCTCTGCGCAAGTGTGTCACAAATACTCCACTGTGTATCTTTCACAGCACAG 960
Qy 999 cctgtcgtatcgtagcgtcctcaccatcgtcgcagatgtctgttctgaactgcgcag 1058
Db 961 CTTGCTGTATCTGAGCTCTCCACATCTGAGAGTCTGATCTGATCTGCTGCTGCTGCTG 1020
Qy 1059 caagcacaagaagcgccctgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgc 1118
Db 1021 CAAGCACAAGAAGCGCGCTGCGGAGCGCTGAGAAAGCTTTTCACAGGCGCACTGGGCAC 1080
Qy 1119 gtgagctgtcgtcgtcgtgtgttgcagcaccactgcgcgcgcgcgcgcgcgcgcgcgc 1178
Db 1081 GTGGCTGCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1140
Qy 1179 gatgagcagacacacgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgc 1238
Db 1141 GATGAGACGACACCCCTGACAGAGCGGAGGAGCAGACAGAGTCCCTGCTGGCATCA 1200
Qy 1239 ccaacaccgagtgccgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgc 1298
Db 1201 CCACACCGAGAGTCCGGGCGCCCAAGGCCAACAGTTGATGGGCGCTGCTGCGCACCGC 1260
Qy 1299 cgtgagcagcattcctcgttctcctcagcctgcgccttcctcatcttgacatcagtg 1358
Db 1261 CGTGGACCGCATTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 1320
Qy 1359 ctccgtgtaggat 1372
Db 1321 CTCCTGTAGGAT 1334

```

```

RESULT 4
AC004573 85095 bp DNA linear INV 30-JAN-1999
LOCUS Drosophila melanogaster, chromosome 2L, region 21C5-21D1, P1 clone
DEFINITION DS07610, complete sequence.
ACCESSION AC004573
VERSION AC004573.1 GI:4204255
KEYWORDS HTG.
SOURCE Fruit fly.
ORGANISM Drosophila melanogaster
Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
Pterygota; Neoptera; Endopterygota; Diptera; Brachycera;
Muscomorpha; Ephydroidea; Drosophilidae; Drosophila.
1 (bases 1 to 85095)
Celinker,S.E., Agbayani,A., Arcaina,T.T., Baxter,E., Blazej,R.G.,
Butenhoff,C., Champe,M., Chavez,C., Chew,M., Ciesiolka,L.,
Doyle,C.M., Farfan,D.E., Galle,R., George,R.A., Harris,N.L.,

```

```

TITLE
JOURNAL Sequencing of Drosophila chromosome 2L, region 21C5-21D1
REFERENCE Unpublished (1998)
AUTHORS
Hoskins,R.A., Houston,K.A., Hummasti,S.R., Karra,K., Kearney,L.,
Kim,E., Lee,B., Lewis,S., Li,P., Lomotan,M.A., Mazda,P.,
Moshrefi,A.R., Moshrefi,M., Nixon,K., Pacleb,J.M., Park,S.,
Pfeiffer,B., Poon,L., Sequeira,A., Setchell,H., Snit,E.,
Svirskas,R.R., Wan,K.H., Weinburg,T., Zhang,R., Zieran,L.L. and
Rubin,G.M.
2 (bases 1 to 85095)
Celinker,S.E., George,R.A., Galle,R., Svirskas,R.R., Hoskins,R.A.,
Agbayani,A., Arcaina,T.T., Baxter,E., Blazej,R.G., Chavez,C.,
Chew,M., Doyle,C.M., Farfan,D.E., Flanagan,J., Houston,K.A.,
Hummasti,S.R., Karra,K., Kearney,L., Kim,S.H., Lee,B.,
Lomotan,M.A., Mak,J., Mazda,P., Mock,M.S., Moshrefi,A.R.,
Moshrefi,M., Nixon,K., Pacleb,J.M., Park,S., Pfeiffer,B., Punch,E.,
Snit,E., Twomey,B., Wan,K.H., Whitehaw,K.R., Yee,A., Zhang,R.,
Zieran,L.L. and Kimmel,B.E.
Direct Submission
TITLE
JOURNAL Submitted (17-APR-1998) Drosophila Genome Center, Lawrence Berkeley
Laboratory, MS 64-121, Berkeley, CA 94720, USA
COMMENT On Jan 30, 1999 this sequence version replaced gi:4164118.
Sequence submitted by:
Berkeley Drosophila Genome Project
Lawrence Berkeley National Laboratory, MS 64-121
Berkeley, CA 94720
For further information about this sequence, including its location
and relationship to other sequences, please visit our sequence
archive web site (http://www.fruitfly.org/sequence/) or send email
to bdg@fruitfly.berkeley.edu.
P1 library location: 80-26.
FEATURES
location/Qualifiers
1..85095
/organism="Drosophila melanogaster"
/strain="y2: cn bw sp"
/db_xref="taxon:7227"
/chromosome="2L"
/map="21C5-21D1"
/clone="P1 DS07610 (D275)"
/clone.lib="P1 library, partial Sau3a in pNS582etclAAd10"
BASE COUNT 22289 a 19671 c 20319 g 22816 t
ORIGIN
Query Match 40.8%; Score 628; DB 3; Length 85095;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 628; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 352 cgaatggaagacgagagcggtgtggaacgclcaacaatagacaacacacgagatc 411
Db 5406 CGATGGAGACAGCAGAGCGCGTGTGGCAACCGTCAACATATGACAAATCAGCAGATC 5465
Qy 412 actttgaatcgcagcagatgtctggaacccccaataacactcttcaaggcaggaagt 471
Db 5466 ACTTTGAAGTCCAGCGAGGCTGTGAGCCGCCCAATACACTCTTCAACGCGCAGAGGT 5525
Qy 472 ggcctgaatgcggaaccccgatgacccctcagacagatgacacattccggtggaatgct 531
Db 5526 GGCTGATGCGCGAAGACCCAGGTGACCTCAAGCAGATGACCACTTCGCGGAGAACGCGCT 5585
Qy 532 ccaagcctgtacacgagcctactcgaactcaaatcagctcgaactgcgcgcgcgcgcgcgcgc 591
Db 5586 CCAAGCGGTATCAACGCGCTACTCGAATCTCAACATGCTCAACTGCGCCCAAGACAGAG 5645
Qy 592 agctgcagatggaatcgcgtccctgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgc 651
Db 5646 AGCTGCAAGTTGAAGATCGGCTCTCTGAGGCTCTGCTCTCTGCGGAGAACGCGCAGC 5705
Qy 652 ggcgaagaagatccctctgacacagacgactggttaagtcacacgagtggaatcgtg 711
Db 5706 GCGAGAGAGATCCCTTGACCAACGACGACTGTTTACGACGCGGATGGGAATGCTG 5765
Qy 712 gactcgaagccactgttcagtcaggaactaagcgtcacatgagtaactcgtgacg 771

```

[illegible]

	Db	12167	AGCTGCAAGTGTGAAGATGGCGCTCCTGGGGCCGTGAAGGTCGTCCTGC CGSAGAAGCGACAG	121967
	Qy	652	gcgagaggagagctcccttgaccacgaagcactggttcacgaacggagatggaaatcgtg	711
	Db	121966	GCGAGAGAGAGATGCCCTTGACCACGACACCACTGGTTCAAGTACCCGAGTGGGAAATCGGTG	121907
	Qy	712	gacctcgagagccaacttggtaagtcagaagcaactacaagccataagagtagtacatctgcg	771
	Db	121906	GACTGCGCAGGCCCATTTGTGTAAGTACGAGACTAAGTACGAGCTATATGAGATTACCTTGACG	121847
	Qy	772	gctcacggcgagctccctcccatgtlacacggcgctlcatctacacacccgcgtcccgatcgtc	831
	Db	121846	GCTCAGCGGGCGCTCCTCCATGTACACGCGCGCATCTAACACCCGCGCTCGATCGTC	121787
	Qy	832	atcctgagcctctcgaagcctcttgctgctgcctcccacatggcggcagagaagatcatgac	891
	Db	121786	ATCCCTGGCCCTCTCAGCCCTTGTGGCTGCCTCCCACTAATGAGGGGCGGGAAGAATCATGTGATC	121727
	Qy	892	aagggcctgcacatcatgatgatcgccgcttcctccatgactcttcgccagactcctgcga	951
	Db	121726	AACGGCCGCTCATCATGTATGTATGCCCGCCCTTCCATCATGTACTTGGCCGACGACTCTTGCCA	121667
	Qy	952	gtgcgtgcacaacatattccactgtg	979
	Db	121666	GTGCTGTCCACAACTATCTCACCTTGCG	121639
RESULT	6			
AC008002/c			182726 bp DNA Linear INV 06-APR-2001	
LOCUS			Drosophila melanogaster, chromosome 2L, region 21D-21E, BAC clone	
DEFINITION			BACR4BE08, complete sequence.	
ACCESSION			AC008002	
VERSION			AC008002.3 GI:13559545	
KEYWORDS			HTG.	
SOURCE			fruit fly.	
ORGANISM			Drosophila melanogaster	
REFERENCE			Eukaryota; Neoptera; Endopterygota; Diptera; Brachycera;	
AUTHORS			Muscomorpha; Ephydroidea; Drosophilidae; Drosophila.	
			1 (bases 1 to 182726)	
			Celniker,S.E., Adams,M.D., Krommler,B., Tyler,D., Wan,K.H.,	
			Holt,R.A., Evans,C.A., Gocayne,J.D., Amanatides,P.G., Brandon,R.C.,	
			Rogers,Y., An,H., Baldwin,D., Banzon,J., Beeson,K.Y., Busan,D.A.,	
			Carlson,J.W., Center,A., Chapple,M., Davenport,L.B., Dietz,S.M.,	
			Dodson,K., Dorsett,V., Doup,E.E., Doyle,C., Dresnek,D., Farfan,D.,	
			Ferreria,S., Frishe,E., Galile,R.F., Gard,N.S., George,R.A.,	
			Gonzalez,M., Houck,J., Hoskins,R.A., Hostin,D., Howland,T.J.,	
			Idegian,C., Jalali,M., Kruse,D., Li,P., Mattei,B., Moshrefi,A.,	
			McIntosh,T.C., Moy,M., Murphy,B., Nelson,C., Nelson,K.A., Nunoo,J.,	
			Pacle,D., Paragas,V., Park,S., Patel,S., Pfeiffer,B.,	
			Phouaneavong,S., Piltman,G.S., Purf,V., Richards,S., Scheeler,F.,	
			Stapleton,M., Strong,R., Svirskas,R., Tector,C., Williams,S.M.,	
			Zaveri,J.S., Smith,H.O., Rubin,G.M. and Venter,J.C. 21E	
TITLE			Sequencing of Drosophila chromosome 2L, region 21D-21E	
JOURNAL			Unpublished	
REFERENCE			2 (bases 1 to 182726)	
AUTHORS			Celniker,S.E., Agbayan,A., Arcaina,T.T., Baxter,E., Blazek,R.G.,	
			Baltenhoff,C., Champe,M., Chavez,C., Chew,M., Ciesiolka,L.,	
			Doyle,C.M., Farfan,D.E., Galile,R., George,R.A., Harris,N.L.,	
			Hoskins,R.A., Houston,K.A., Hummasti,S.R., Katra,K., Kearney,L.,	
			Khm,E., Lee,B., Lewis,S., Li,P., Lomotan,M.A., Mazda,P.,	
			Moshrefi,A.R., Moshrefi,M., Nixon,K., Pacleb,J.M., Park,S.,	
			Pfeiffer,B., Poon,L., Sequeira,A., Sethi,H., Snir,E.,	
			Svirskas,R.R., Wan,K.H., Weinburg,T., Zhang,R., Zieran,L.L. and	
			Rubin,G.M.	
TITLE			Direct Submission	
JOURNAL			Submitted (09-JUL-1999) Drosophila genome center, Lawrence Berkeley	
REFERENCE			Laboratory, MS 64-121, Berkeley, CA 94720, USA	
AUTHORS			On Apr 6, 2001 this sequence version replaced gi:16532012.	
COMMENT			Sequence submitted by:	
			Lawrence Drosophila Genome Project	
			Berkeley Berkeley National Laboratory, MS 64-121	


```
gene
<3618..>39613
/gene="PLC21C"
/note="CG4574"
/map="21B7-21C1"
/db_xref="FLYBASE:FBan0004574"
/db_xref="FLYBASE:FBgn0004611"
join(11383..11475,19033..19179,27475..27612,28770..28903,
29373..29558,29704..29901,30371..30553,31293..31473,
32562..32796,34065..34210,35933..36146,36220..37083,
37396..37843,37909..38301,38363..38507,38569..38802)
/gene="PLC21C"
/note="PLC21C gene product"
/codon_start=1
/db_xref="FLYBASE:FBan0004574"
/db_xref="FLYBASE:FBgn0004611"
/protein_id="AAF51507.1"
/db_xref="GI:7296215"
/translation="MMSAGTYITSAVSVEPOLADGEKFIKMDSDSGTGTPTVTRVD
AKGFLVWDONNELDLDIATIRDTQYAKRPDKNLRLQVLTGLPDQTEKTVT
VCHGSDVFNMTFVNEFCSTRDIAQLMTDGLIKLAVSLAQNGSAIMFLOKATKLCIQ
VKSGRIIPYKNIIKLFPAONKEDRKVEKALDYTLPSGVDSISVSKFOEDPYNLK
YLTORSEVERLDSIVGNSKRCMSIALVEFLNKQRDPLNELIYPANARAKEL
IOYEPKNNAKCGSLDGLFLRYLMGDNPTAKPKDLCDMDQPMHPTFNSHN
TTLTGHLTKSSVEVLYROCLLAGCQVLEDFMNGRTEPEVIVHGYTPEIPAKOV
EALAESAFKTSSEVPLISFENHCNPROQAKIANYCEIFGDMLLDPDLSHPLEPMD
LPPAPMLRRKRIIKNKKHHHHHHHHKPPAOVGPANNNKLTANSVDARAAQVG
LPSHEDGVSTANGDVAATGTGTAGTAGHAPLOOIROSSKDSGSDSSS
EESLDPNTPTNLPSCGNEPPEKAKETAGETASALVNYVOPTHFSSPNAEKKNKY
EKSSPERKATTLKRPLEPNYKHOISRYRPACTRFDSSNFMQLFMNNGCOLVA
LNFOTLDLMQMLGFEYNASGYLKLPEFMRSRDRDLPRESTVDGIAGTASIT
VLSGFLTDKRANTEVEVDVYGLPADTVRKFKETKVRNDGNMPLDEEPFKVYL
PELASIRIAVEGKGLIGHRYLVGLGPKYKHYMLRSEVGPILALAEFLCYVVD
IYDDLSNKAALANAIKYOSELEKRDIOLVTLDEALASADELSKSCGKELR
PVESLTSKPHRPSISAAAMSVTVDDRGREGDSISVAPSIQHOHSQST
SITROVSSQEDVDIVLAEPLKILDKSVKEREKMEKRIKHKHDEKIKIADOK
SEPLECKRKPATIKLYRKLSKSLNCSLSPHSEPCVEIPACPLDGLDSSSEBAAD
GEDLAGSSSLDGRTOESRLRSACREYSQYRSLQRYEATYSAEKVLTQSTQYT
KOLKALDFRDTDELQKODSVARNALAKATVIRDEDELIRMKREVAASSVEGVAR
VRLKQTFDRDTDELQKODSVARNALAKATVIRDEDELIRMKREVAASSVEGVAR
HHHGCGSSSSALSGNNLTFLNDAGNAGSHSATSPSKSHNSIAAAEMKT"
complement(11712..12601)
/gene="CG11912"
/product="CT37080"
/db_xref="FLYBASE:FBan0011912"
/db_xref="FLYBASE:FBgn0031248"
/complement(<11712..>12601)
/gene="CG11912"
/map="21B7-21B8"
/db_xref="FLYBASE:FBan0011912"
/db_xref="FLYBASE:FBgn0031248"
/complement(11773..12588)
/gene="CG11912"
/note="CG11912 gene product"
/codon_start=1
/db_xref="FLYBASE:FBan0011912"
/db_xref="FLYBASE:FBgn0031248"
/protein_id="AAF51509.1"
/db_xref="GI:7296217"
/translation="MKQFAVIFALASVATISVPDGPPEGRITNGYBAKGEAPYI
VSIQTSSNHFCAAGSLDDEVITYTAHCHCTYNGQAVAAHSTDEBNQIKETNAQ
YVITHENYGGVGNPDIGLTLLEKEDAFDLNAARQDSNPNVSAVLSKTFQGSIDYL
YWGGRNSGLPLNLOKDAIIVDYNECKAALPNSNLSAETNVCITHTPKRADSSCGD
SGGPIVSSSSRAELIGIVSMGVTPLCSTTTPSVYTSVSSFLPMWIDENKKA"
complement(13648..14580)
/gene="CG11911"
/product="CT37076"
/db_xref="FLYBASE:FBan0011911"
/db_xref="FLYBASE:FBgn0031249"
/complement(<13648..>14580)
/gene="CG11911"
/map="21B7-21B8"
/db_xref="FLYBASE:FBan0011911"
/db_xref="FLYBASE:FBgn0031249"
/gene
```

```
CDS
complement(13743..14576)
/gene="CG11911"
/note="CG11911 gene product"
/codon_start=1
/db_xref="FLYBASE:FBan0011911"
/db_xref="FLYBASE:FBgn0031249"
/protein_id="AAF51508.1"
/db_xref="GI:7296216"
/translation="MKLIIVTIVIALVAAGAKLSDKLAKVPSFANGFVINGTEAE
PHSAPIVALANVYLKSHIGCTGLINKMIVTAHNCISEVPGMSIAGLHTRAEVDE
LTOORQVDEGRVHERKYTGCGPYDIALHLVNSFTFENWVOPATLPSRQVGEFTL
YMGQPKSYTESGAKTLQVTTQIALLNECKEELPESAPIASNSICSSLDQSKACH
GDSGPELVVEFTNAPSELIGIVSMGVTPLGLNMPSTYTKVASIYDWTITNDSATYKL
F"
complement(join(49281..49459,49730..50200,50254..50554,
50786..>51230))
/gene="CG11907"
/product="CT33136"
/db_xref="FLYBASE:FBan0011907"
/db_xref="FLYBASE:FBgn0031250"
/complement(<49281..>51230)
/gene="CG11907"
/map="21C1-21C1"
/db_xref="FLYBASE:FBan0011907"
/db_xref="FLYBASE:FBgn0031250"
/complement(join(49281..49459,49730..50200,50254..50554,
50786..51049))
/gene="CG11907"
/note="CG11907 gene product"
/codon_start=1
/db_xref="FLYBASE:FBan0011907"
/db_xref="FLYBASE:FBgn0031250"
/protein_id="AAF51506.1"
/db_xref="GI:7296214"
/translation="MGTSDTRIVDALDEDEDELITGESFVPAPRPDHHDEOCIM
ERQADEVYVSNESPESRGFTYVLYELDGLIGMTVPMNEPTVTEDEWKYKRRASINT
DLEEDTFLQKSFCDALATATISGTFLILALIGHNHSLETKMGLIMTLIFGV
TTFGEINIDTMOEFOFLTLLIIVLNTISAATMSGALYGVAGLEPSEFTIYVSGQA
LGCIITALAFIIVLAFDTPGNTTAFEFYVGVLLILCYCYVILAKRPFYVLEGG
DKYIVRAVPSHNRRNSAGLEPLEPILROMVSKIYLHAISLALVTTTLYSVPAATVL
MOSEYGSVMTAMVMWFALSNGYFTNILLIMPKRVKQREKELASIMAAALSCMAV
GSLSLVPOYL"
join(68682..69037,75067..75450,76257..76302,76864..77332,
77259..77995)
/gene="al"
/note="Nucleotide sequence of the Celera sequence differs
from the published sequence for this transcript."
/product="CT12963"
/db_xref="FLYBASE:FBan0003935"
/db_xref="FLYBASE:FBgn0000061"
<68682..>77995
/gene="al"
/note="CG3935"
/map="21C1-21C2"
gene
mRNA
Query Match 40.8% Score 628 DB 3: Length 302473:
Best Local Similarity 100.0% Pred. No. 0:
Matches 628; Conservative 0; Mismatches 0; Indels 0; Gaps 0:
QY 352 cgatgaggaagaagagcgctgtgcaacccgtcccaatatatacaaacatcagcgatc 411
|||||
Db 236373 CGATGGAAGACGACGAGCGCGTGTGCAACCGCTCAACAAATATACAAACATCAGCAGATC 236432
QY 412 accttgagtcagcgaggtctgagcccccacaatcacacatctctcaacgcgagcgaagt 471
|||||
Db 236433 ACTTTGAAGTCACGACGAGCTGTGACCCGCCCAATATCACTCTTCAAGCGCAGCAAGGT 236492
QY 472 gacctgagtcagcgaacccaggtgacccctcagcagaatgagtgcaacttcggtgagtgacct 531
|||||
Db 236493 GGCCTGATGCGCGAAGCCACGAGTGACCTCAGCCACAGATGCCACTGCCGTGATGCCCT 236552
QY 532 ccagcgctgacacgagctactcgaaactcaaatgtcaactcggcccgacgaagcag 591
|||||
```

Db 236553 CCAGCCGCTGTACAGCGGCTACTGCGAATCAATGCTCAACTGCGCCCAAGACAGACG 236612
QY 592 agcggcaagtgaagatcgcgtcctgggagctgaaagtgcctcgcgggagaacgacg 651
Db 236613 AGCTGCAAGTTGAAGATCGGCTCTGGGGCTGAAGGTCGTCTCCGCGAAGACGACG 236672
QY 652 gcggaggaagatgctccctgacacgacgacacgtgctcaatcaccgagtgagaaatcgtg 711
Db 236673 GCGAGAGAGAGTCCCTTGACACGACGACACTGGTTCACTACCGGAGTGGGAAATCGTG 236732
QY 712 gactcgcagacccacttctgtaagtaactactacgactacgtatagatcactctgacg 771
Db 236733 GACTCGGAGAGCCACTTGTGTAGTACAGACTACTACGGCTCATGAGTACACTCTGACG 236792
QY 772 gctagaggagcgtcctcctcatgtatagagccgctcatctacaccccgctcctgcatgctc 831
Db 236793 GCTAGAGGGGCGCTCTCCATGTATACAGCGCGCTCATACACACCGCGCTCGATCGTC 236852
QY 832 atcctgagccctctcagcctctgctgctccacatagggcggagaaatcatgac 891
Db 236853 ATCTGGCGCTCTCAGCGCTTCTGGCTGCTGCCACATGGGCGGCGAAGATCATGATC 236912
QY 892 aagcgctgcctcatcgtatgcgcgctcctcctcatgtactctgcgcagctcctgcca 951
Db 236913 AACGCGCTGCTCATCATCATGATGATGCGCGCTTCTCATGTACTTTCGCCACGCTCGCA 236972
QY 952 gtgctgtccacaatactcactcgtg 979
Db 236973 GTGCTGTCCAACTACTCCACTTGTG 237000

RESULT 8
AY061622 1646 bp mRNA linear INV 08-NOV-2001
LOCUS Drosophila melanogaster SD09326 full length cDNA.
DEFINITION AY061622
ACCESSION AY061622.1 GI:16769901
VERSION FLI.CDNA.
KEYWORDS fruit fly.
SOURCE Drosophila melanogaster
ORGANISM Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
Pterygota; Neoptera; Endopterygota; Diptera; Brachycera;
Muscomorpha; Ephydroidea; Drosophilidae; Drosophila.
1 (bases 1 to 1646)

REFERENCE
AUTHORS Stapleton, M., Brokstein, P., Hong, L., Abmayan, A., Carlson, J.,
Champe, M., Chavez, C., Dorsett, V., Farfan, D., Frise, E., George, R.,
Gonzalez, M., Guarin, H., Li, P., Liao, G., Miranda, A., Mungall, C. J.,
Nunoo, J., Pacle, J., Paragas, V., Park, S., Phoumenavong, S., Wan, K.,
Yu, C., Lewis, S. E., Rubin, G. M. and Celisner, S.
Direct Submission

TITLE Submitted (30-OCT-2001) Berkeley Drosophila Genome Project,
JOURNAL Lawrence Berkeley National Laboratory, One Cyclotron Road,
Berkeley, CA 94720, USA

COMMENT Sequence submitted by:
Berkeley Drosophila Genome Project
Lawrence Berkeley National Laboratory
Berkeley, CA 94720

This clone was sequenced as part of a high-throughput process to
sequence clones from Drosophila Gene Collection 1 (Rubin et al.,
Science 2000). The sequence has been submitted to integrity checks
for sequence accuracy, presence of a polyA tail and contiguity
within 100 kb in the genome. Thus we believe the sequence to
reflect accurately this particular cDNA clone. However, there are
artifacts associated with the generation of cDNA clones that may
have not been detected in our initial analyses such as internal
priming, priming from contaminating genomic DNA, retained introns
due to reverse transcription of unsplined precursor RNAs, and
reverse transcriptase errors that result in single base changes.
For further information about this sequence, including its location
and relationship to other sequences, please visit our web site
(http://fruitfly.berkeley.edu) or send email to
cnae@fruitfly.berkeley.edu.

FEATURES

Location/Qualifiers

source 1. .1646
/organism="Drosophila melanogaster"
/db_xref="taxon:7227"
/map="21C5-21C5"
gene 1. .1646
/gene="naCRbeta-21C"
/note="alignment with genomic scaffold AE003589"
/db_xref="FLYBASE:FBgn0031261"
CDS 37. .1020
/gene="naCRbeta-21C"
/note="longest ORF"
/codon_start=1
/db_xref="FLYBASE:FBgn0031261"
/product="SD09326p"
/protein_id="AA29170.1"
/db_xref="GI:16769902"
/translation="MTTPKIKAPVSGPLPLLLMGLMGLTSLVPATATADPK
NANVKGIDRLHAGLFTNDSYDVPFPGFTPTSLGMVYVYIDDELNGKLTWMLN
LRWDERWVQPSOYDNTIQITLKSSEWTPQITLNGDGLMAETQYTLSDGHR
WMPAYVTATACELNMLNMPDRKSCRLKIGSMGLKVLPLPENGARDESIDHDLOSP
EMETVSRATFVQDYCYGYEYTLTQNRSSMTYVITYPACTIVLALSAFWLPPHM
GSEKIMINGLLITVIAFLMTFQQLPLVLSNNTPLVGECSPLPFHOFSSBN"
BASE COUNT 390 a 506 c 428 g 322 t
ORIGIN
Query Match 32.9%; Score 507; DB 3; Length 1646;
Best Local Similarity 99.1%; Pred. No. 1e-269;
Matches 957; Conservative 0; Mismatches 9; Indels 0; Gaps 0;
QY 14 gtatccgaacaaagcgcgctgaaacatgacgacgactcccaagataaagcaccag 73
Db 8 GTATATCGGAACAAAGCGCGCTGAAACATGACGACGACTCCCAAGATAAGCACCAG 67
QY 74 ttccggtcctgacgctgacactactgctgcaaatgtaatgaggaatgctcttaagggc 133
Db 68 TTTCCGGTCTCGAGCTGCACTACTGCTGCAAAATGATGAGGATGCTTCTTATGGGGC 127
QY 134 tgaattccggtgacgagggcactctgcaacggcgagcccaagaacgcaatgtaagggc 193
Db 128 TGACTTCCGTGCGAGGCGCACTGCTCCACCGCGGCCCAAGAACGCAATGTCAAGGGC 187
QY 194 tggatgcctcgaacggcgctgttcaagacactcgaacgagatgctgacggctgttcc 253
Db 188 TGGATCGCTTCACGCGGCGCTGTTTCACAGAACTACGACAGAGATGTCAGCGGTTC 247
QY 254 aaggaaccccccaagacgctgctccctggaatgctgctcaactacatagacatcgaagat 313
Db 248 AAGGAACCCCAAGCAAGTGTCTCGGAATGATGTGACCTTACATGACATCGACAGT 307
QY 314 tgaacggcaagctgacacccactgctgctgtaattccgattgagagagagagcgcg 373
Db 308 TGAACGGCAAGCTGACCACTGCTGCTGAAATCTCCGATGGAAGACGAGAGCGCG 367
QY 374 ttgtgcaacgctcaaatatgacacatcgaacgacgacttgaattgaatccgagggct 433
Db 368 TGTGCAACCTCATCATATATGCAACATGACGACGAGATCATCTTGAATCCAGCAGTCT 427
QY 434 ggaaccccccaatcacactcttcaacggcgaagatgctgctgtaagtcggaacccag 493
Db 428 GGACCCCCCAAAATCACACTCTTCAACGGCGACGAAGCGGCGCTGATGGCCGAACCCAG 487
QY 494 tgaccctcagcgaagatggaacttcggtgagatgctcgaacggctgtaacggcctact 553
Db 488 TGACCTCTAGCAAGATGCGCATGATGATGATGATGATGATGATGATGATGATGATG 547
QY 554 gcgaacccaacatgctcaactgacgcccccaagcgaagagctggaagtgaagtcggct 613
Db 548 GCGAATCTCAATCTCAATGCGCCCAAGACGACGAGAGCTGCAACTGAAAGTCGGCT 607
QY 614 cctggagcctgaaagtcgtcctgctcggaagacgacgagcgaagagagagtcctgacc 673
Db 608 CTTGGGCGCTGAAGTGTCTGCGGGAAGACGGCGACGAGAGAGAGTCTTGTACC 667

* sequencing reads that have not been assembled into
* contigs. Runs of N are used to separate the reads
* and the order in which they appear is completely
* arbitrary. Low-pass sequence sampling is useful for
* identifying clones that may be gene-rich and allows
* overlap relationships among clones to be deduced.
* However, it should not be assumed that this clone
* will be sequenced to completion. In the event that
* the record is updated, the accession number will
* be preserved.

1 714: contig of 714 bp in length
715 814: gap of 100 bp
815 1512: contig of 698 bp in length
1513 1612: gap of 100 bp
1613 2314: contig of 702 bp in length
2315 2414: gap of 100 bp
2415 3125: contig of 711 bp in length
3126 3225: gap of 100 bp
3226 3934: contig of 709 bp in length
3935 4034: gap of 100 bp
4035 4748: contig of 714 bp in length
4749 4848: gap of 100 bp
4849 5553: contig of 705 bp in length
5554 5653: gap of 100 bp
5654 6354: contig of 701 bp in length
6355 6454: gap of 100 bp
6455 7167: contig of 713 bp in length
7168 7267: gap of 100 bp
7268 7979: contig of 712 bp in length
7980 8079: gap of 100 bp
8080 8798: contig of 719 bp in length
8799 8898: gap of 100 bp
8899 9592: contig of 694 bp in length
9593 9692: gap of 100 bp
9693 10393: contig of 701 bp in length
10394 10493: gap of 100 bp
10494 11210: contig of 717 bp in length
11211 11310: gap of 100 bp
11311 12030: contig of 720 bp in length
12031 12130: gap of 100 bp
12131 12843: contig of 713 bp in length
12844 12943: gap of 100 bp
12944 13662: contig of 719 bp in length
13663 13762: gap of 100 bp
13763 14462: contig of 700 bp in length
14463 14562: gap of 100 bp
14563 15265: contig of 703 bp in length
15266 15365: gap of 100 bp
15366 16077: contig of 712 bp in length
16078 16177: gap of 100 bp
16178 16855: contig of 678 bp in length
16856 16955: gap of 100 bp
16956 17667: contig of 712 bp in length
17668 17767: gap of 100 bp
17768 18464: contig of 697 bp in length
18465 18564: gap of 100 bp
18565 19271: contig of 707 bp in length
19272 19371: gap of 100 bp
19372 20088: contig of 718 bp in length
20090 20189: gap of 100 bp
20190 20914: contig of 725 bp in length
20915 21014: gap of 100 bp
21015 21716: contig of 702 bp in length
21717 21816: gap of 100 bp
21817 22532: contig of 716 bp in length
22533 22632: gap of 100 bp
22633 23339: contig of 707 bp in length
23340 23439: gap of 100 bp
23440 24142: contig of 703 bp in length
24143 24242: gap of 100 bp
24243 24951: contig of 709 bp in length
24952 25051: gap of 100 bp
25052 25762: contig of 711 bp in length

25763 25862: gap of 100 bp
25863 26576: contig of 714 bp in length
26577 26766: gap of 100 bp
26767 27379: contig of 703 bp in length
27380 27479: gap of 100 bp
27480 28196: contig of 717 bp in length
28197 28296: gap of 100 bp
28297 29018: contig of 722 bp in length
29019 29118: gap of 100 bp
29119 29843: contig of 725 bp in length
29844 29943: gap of 100 bp
29944 30664: contig of 721 bp in length
30665 30764: gap of 100 bp
30765 31484: contig of 720 bp in length
31485 31584: gap of 100 bp
31585 32271: contig of 687 bp in length
32272 32371: gap of 100 bp
32372 33070: contig of 699 bp in length
33071 33170: gap of 100 bp
33171 33884: contig of 714 bp in length
33885 33984: gap of 100 bp
33985 34691: contig of 707 bp in length
34692 34791: gap of 100 bp
34792 35505: contig of 714 bp in length
35506 35605: gap of 100 bp
35606 36310: contig of 705 bp in length
36311 36410: gap of 100 bp
36411 37132: contig of 722 bp in length
37133 37232: gap of 100 bp
37233 37917: contig of 685 bp in length
37918 38017: gap of 100 bp
38018 38725: contig of 708 bp in length
38726 38825: gap of 100 bp
38826 39545: contig of 720 bp in length
39546 39645: gap of 100 bp
39646 40358: contig of 713 bp in length
40359 40458: gap of 100 bp
40459 41171: contig of 713 bp in length
41172 41271: gap of 100 bp
41272 41982: contig of 711 bp in length
41983 42082: gap of 100 bp
42083 42767: contig of 685 bp in length
42768 42867: gap of 100 bp
42868 43571: contig of 704 bp in length
43572 43671: gap of 100 bp
43672 44384: contig of 713 bp in length
44385 44484: gap of 100 bp
44485 45190: contig of 706 bp in length
45191 45290: gap of 100 bp
45291 45987: contig of 697 bp in length
45988 46087: gap of 100 bp
46088 46800: contig of 713 bp in length
46801 46900: gap of 100 bp
46901 47620: contig of 720 bp in length
47621 47720: gap of 100 bp
47721 48427: contig of 707 bp in length
48428 48527: gap of 100 bp
48528 49260: contig of 733 bp in length
49261 49360: gap of 100 bp
49361 50055: contig of 695 bp in length
50056 50155: gap of 100 bp
50156 50862: contig of 707 bp in length
50863 50962: gap of 100 bp
50963 51643: contig of 681 bp in length
51644 51743: gap of 100 bp
51745 52444: contig of 701 bp in length
52445 52544: gap of 100 bp
52545 53233: contig of 689 bp in length
53234 53333: gap of 100 bp
53334 54034: contig of 701 bp in length
54035 54134: gap of 100 bp
54135 54841: contig of 707 bp in length
54842 54941: gap of 100 bp

* 54942 55656: contig of 715 bp in length
* 55657 55756: gap of 100 bp

Query Match 1.6%; Score 24; DB 2; length 74135;
Best Local Similarity 100.0%; Pred. No. 0.24;
Matches 24; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1516 tctgcttcaaaaaaaaaaaaaa 1539
|||||
Db 58316 TCTGCTTACAAAAAAA 58339

RESULT 11
AC096600 77575 bp DNA linear HTG 20-DEC-2001
LOCUS Rattus norvegicus clone CH230-111A10, *** SEQUENCING IN PROGRESS
DEFINITION *** 40 unordered pieces.
ACCESSION AC096600 GI:17944006
VERSION AC096600.2
KEYWORDS HTG; HTGS_PHASE1.
ORCE Norway rat.
ORGANISM Rattus norvegicus
Eukaryota; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
Rattus.
1 (bases 1 to 77575)

REFERENCE
AUTHORS
Muzny,D.M., Adams,C., Adio-Oduola,B., All-rosman,F.R., Allen,C.,
Alspbrooks,S.L., Amaralunge,H.C., Are,J.R., Banks,T., Barbaria,J.,
Benton,J., Bimago,K., Blankenburg,K., Bonnin,D., Bouck,J.,
Bowie,S., Brieva,M., Brown,E., Brown,M., Bryant,N.P., Buhay,C.,
Burck,P., Burkett,C., Burrell,K.L., Byrd,N.C., Carron,T.F.,
Carter,M., Cavazos,S.R., Chacko,J., Chavez,D., Chen,G., Chen,R.,
Chen,Z., Chowdhury,I., Christopoulos,C., Cleveland,C.D., Cox,C.,
Coyle,M.D., Dathorne,S.R., David,R., Davila,M.L., Davis,C.,
Davy-Carroll,L., Dederich,D.A., Delaney,K.R., Delgado,O.,
Dem,A.L., Ding,Y., Dinh,H.H., Douthwaite,K.J., Draper,H.,
Dugan-Rocha,S., Durbin,K.J., Earnhart,C., Edgar,D., Edwards,C.C.,
Elhaj,C., Escotto,M., Falls,T., Ferraguto,D., Flagg,N., Ford,J.,
Foster,P., Frintz,P., Gabisi,A., Gao,J., Garcia,A., Garner,T.,
Garza,N., Gill,R., Gorrell,J.H., Guevara,M., Gunaratne,P., Hale,S.,
Hamilton,K., Harris,C., Harris,K., Hart,M., Havlak,P., Hayes,A.,
Hernandez,J., Hernandez,O., Hodgson,A., Hogue,M., Holloway,C.,
Hollins,B., Homs,F., Howard,S., Huber,J., Hulyk,S., Hume,J.,
Jackson,L.E., Jacobson,B., Jia,Y., Johnson,R., Jollivet,S.,
Joudah,S., Karlsson,E., Kelly,S., Khan,U., King,L., Korvah,J.,
Kovar,C., Kratovic,J., Kureshi,A., Landry,N., Leal,B., Lewis,L.C.,
Lewis,L., Li,J., Li,Z., Lichtarge,O., Lieu,C., Liu,J., Liu,W.,
Louiased,H., Lozado,R.J., Lu,X., Lucier,A., Lucier,R., Luna,R.,
Ma,J., Maheshwari,M., Mapua,P., Martin,R., Martindale,A.,
Martinez,E., Massey,E., Mawhinney,E., McLeod,M.P., Meador,M.,
Mel,G., Metzker,M., Miner,G., Miner,Z., Mitchell,T., Mohabbat,K.,
Morgan,A., Morris,S., Moser,M., Neal,D., Newton,J., Newton,N.,
Nguyen,A., Nguyen,N., Nguyen,N., Nickerson,E., Nwokwoko,S.,
Ogulu,M., Okunolu,G., Oragunye,N., Oviedo,R., Pace,A., Payton,B.,
Peery,J., Perez,L., Peters,L., Pickens,R., Primus,E., Pu,L.L.,
Quiles,M., Ren,Y., Rives,M., Rojas,A., Rojucokan,I., Rolfe,M.,
Ruiz,S., Severy,G., Scherer,S., Scott,G., Shen,H., Shoshchari,N.,
Sisson,I., Sodergren,E., Sonaike,T., Sparks,A., Stanley,H.,
Stone,H., Sutton,A., Svatek,A., Taber,P., Tamerisa,A., Tamerisa,K.,
Tang,H., Tansey,J., Taylor,C., Taylor,T., Telford,B., Thomas,N.R.,
Thomas,S., Usmani,K., Vasquez,L., Vera,V., Villalob,D., Vinson,R.,
Wall,R., Wang,S., Ward-Moore,S., Warren,R., Washington,C.,
Washington,S., Williams,G., Williamson,A., Wleczky,R., Wooden,S.,
Worley,K., Wu,C., Wu,Y., Wu,Y.F., Zhou,J., Zorrilla,S., Nelson,D.,
Wenstock,G. and Gibbs,R.

TITLE
JOURNAL
REFERENCE
2 (bases 1 to 77575)

TITLE
AUTHORS
JOURNAL
Submitted (19-SEP-2001) Human Genome Sequencing Center, Department
of Molecular and Human Genetics, Baylor College of Medicine, One
Baylor Plaza, Houston, TX 77030, USA

COMMENT

On Dec 20, 2001 this sequence version replaced gi:15667959.

----- Genome Center
Center: Baylor College of Medicine
Center code: BCM

Web site: <http://www.hgsc.bcm.tmc.edu/>
Contact: hgsc-help@bcm.tmc.edu

----- Project Information
Center project name: GEXD
Center clone name: CH230-111A10
----- Summary Statistics

Assembly program: Phrap; version 0.990329first call to
findPhraplist

Consensus quality: 66157 bases at least Q40
Consensus quality: 70468 bases at least Q30
Consensus quality: 74107 bases at least Q20
Estimated insert size: 57523; sum-of-coverage estimation
Quality coverage: 0x in Q20 bases; agarose-gel estimation
Quality coverage: 0.7x in Q20 bases; sum-of-coverage estimation

* NOTE: Estimated insert size may differ from sequence length
(see http://www.hgsc.bcm.tmc.edu/docs/genbank_draft_data.html).

* NOTE: This is a 'working draft' sequence. It currently
* consists of 40 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.

1	3032:	contig of 3032 bp in length
3033	3132:	gap of unknown length
3133	4681:	contig of 1549 bp in length
4682	4781:	gap of unknown length
4782	7899:	contig of 3118 bp in length
7900	7999:	gap of unknown length
8000	9594:	contig of 1595 bp in length
9595	9694:	gap of unknown length
9695	13290:	contig of 3596 bp in length
13291	13390:	gap of unknown length
13391	17444:	contig of 4054 bp in length
17445	17544:	gap of unknown length
17545	20067:	contig of 2533 bp in length
20068	20167:	gap of unknown length
20168	23171:	contig of 3004 bp in length
23172	23271:	gap of unknown length
23272	25324:	contig of 2053 bp in length
25325	25424:	gap of unknown length
25425	28366:	contig of 2942 bp in length
28367	28466:	gap of unknown length
28467	31210:	contig of 2744 bp in length
31211	31310:	gap of unknown length
31311	33987:	contig of 2677 bp in length
33988	34087:	gap of unknown length
34088	36302:	contig of 2215 bp in length
36303	36402:	gap of unknown length
36403	36867:	contig of 1965 bp in length
36868	38487:	gap of unknown length
38488	40703:	contig of 2216 bp in length
40704	40803:	gap of unknown length
40804	42187:	contig of 1364 bp in length
42188	42287:	gap of unknown length
42288	43715:	contig of 1428 bp in length
43716	43815:	gap of unknown length
43816	45184:	contig of 1369 bp in length
45185	45284:	gap of unknown length
45285	47325:	contig of 2041 bp in length
47326	47425:	gap of unknown length
47426	49810:	contig of 2365 bp in length
49811	49910:	gap of unknown length
49911	51574:	contig of 1664 bp in length
51575	51674:	gap of unknown length
51675	52725:	contig of 1051 bp in length
52726	52825:	gap of unknown length


```

* 52826 54080: contig of 1255 bp in length
* 54081 54180: gap of unknown length
* 54181 55801: contig of 1621 bp in length
* 55802 55901: gap of unknown length
* 55902 56988: contig of 1087 bp in length
* 56989 57088: gap of unknown length
* 57089 58770: contig of 1682 bp in length
* 58771 58870: gap of unknown length
* 58871 60598: contig of 1728 bp in length
* 60599 60698: gap of unknown length
* 60699 61734: contig of 1036 bp in length
* 61735 61834: gap of unknown length
* 61835 62911: contig of 1077 bp in length
* 62912 63011: gap of unknown length
* 63012 64026: contig of 1015 bp in length
* 64027 64126: gap of unknown length
* 64127 65264: contig of 1138 bp in length
* 65265 65364: gap of unknown length
* 65365 66665: contig of 1301 bp in length
* 66666 67844: contig of 1079 bp in length
* 67845 67944: gap of unknown length
* 67945 69145: contig of 1201 bp in length
* 69146 69245: gap of unknown length
* 69246 70292: contig of 1047 bp in length
* 70293 70393: gap of unknown length
* 70393 71450: contig of 1058 bp in length
* 71451 71550: gap of unknown length
* 71551 72603: contig of 1053 bp in length
* 72604 72704: gap of unknown length
* 72704 74424: contig of 1721 bp in length
* 74425 74524: gap of unknown length
* 74525 75961: contig of 1437 bp in length
* 75962 76061: gap of unknown length
* 76062 77575: contig of 1514 bp in length.

```

Location/Qualifiers
1. 77575
/organism="Rattus norvegicus"
/db_xref="taxon:10116"
/clone="CH230-111A10"

BASE COUNT 19126 a 16992 c 16937 t 20577 3943 others

ORIGIN

```

Query Match 1.6%: Score 24; DB 2; Length 77575;
Best Local Similarity 100.0%; Pred. No. 0.24;
Matches 24; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

1516 tctgcttcaaaaaaaaaa 1539
|||||
Db 52890 TCTGCTTCAAAAAAAAAA 52913

RESULT 12
AC090939/c 161361 bp DNA linear PRI 20-MAR-2001
LOCUS Homo sapiens chromosome 3 clone RP11-127114 map 3p, complete
DEFINITION sequence.
ACCESSION AC090939 AC026051
VERSION AC090939.1 GI:13384343
KEYWORDS HTG.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 161361)
AUTHORS Song,L., Bao,J., Bao,Q., Bao,W., Bian,X., Cao,T., Chen,C., Chen,J.,
Ding,H., Dong,W., Fan,H., Feng,X., Gong,J., Guan,Q., Gu,X., Guo,D.,
Guo,Z., He,L., Hu,S., Huang,F., Jin,Y., Kang,N., Li,C., Li,C.,
Li,F., Li,G., Li,J., Li,L., Li,S., Li,T., Liu,Y., Liu,N., Liu,B.,
Liu,Y., Li,W., Li,W., Li,Y., Luo,J., Niu,Y., Qi,O., Qi,X., Song,L.,
Song,S., Sun,M., Sun,W., Sun,Y., Tan,X., Tao,R., Wang,H., Wang,J.,
Wang,J., Wang,J., Wang,L., Wang,L., Wang,R., Wang,X., Wang,X.,

Wang,Y., Wu,D., Wu,Q., Xie,F., Xuan,Z., Xue,Y., Yan,C., Yang,X.,
Yu,B., Zeng,Y., Zhang,G., Zhang,H., Zhang,H., Zhang,L., Zhang,M.,
Zhang,X., Zhang,X., Zhang,Y., Zhang,Y., Zhang,Z., Zhu,B., Zhu,N.,
Yu,J. and Yang,H.
Chromosome 3p genomic sequence
JOURNAL
REFERENCE 2 (bases 1 to 161361)
AUTHORS

Unpublished
Song,L., Bao,J., Bao,Q., Bao,W., Bian,X., Cao,T., Chen,C., Chen,J.,
Ding,H., Dong,W., Fan,H., Feng,X., Gong,J., Guan,Q., Gu,X., Guo,D.,
Guo,Z., He,L., Hu,S., Huang,F., Jin,Y., Kang,N., Li,C., Li,C.,
Li,F., Li,G., Li,J., Li,L., Li,S., Li,T., Liu,Y., Liu,N., Liu,B.,
Liu,Y., Li,W., Li,W., Li,Y., Luo,J., Niu,Y., Qi,O., Qi,X., Song,L.,
Song,S., Sun,M., Sun,W., Sun,Y., Tan,X., Tao,R., Wang,H., Wang,J.,
Wang,J., Wang,J., Wang,L., Wang,L., Wang,R., Wang,X., Wang,X.,
Wang,Y., Wu,D., Wu,Q., Xie,F., Xuan,Z., Xue,Y., Yan,C., Yang,X.,
Yu,B., Zeng,Y., Zhang,G., Zhang,H., Zhang,H., Zhang,L., Zhang,M.,
Zhang,X., Zhang,X., Zhang,Y., Zhang,Y., Zhang,Z., Zhu,B., Zhu,N.,
Yu,J. and Yang,H.
Direct Submission
Submitted (20-MAR-2001) Human Genomic Center, Institute of
Genetics, Chinese Academy of Sciences, Datun Road, Beijing, Beijing
100101, P.R.China
On Mar 20, 2001 this sequence version replaced gi:9958159.
-----Genome Center
Center:Beijing Center
Center code:Beijing
Website:http://hgsc.igtp.ac.cn
http://www.genomics.org.cn
Contact:hgsc@igtp.ac.cn
-----Project Information
Center project name:1% project
Center clone name: RP11-127114
-----Summary Statistics
Sequencing vector: pUC18; 100% of reads
Chemistry: Dye-terminator; EP 5% of reads
Chemistry: Dye-terminator Big Dye; 45% of reads
Assembly program: Phrap; version 0.990329
Consensus quality: 163998 bases at least 400
Consensus quality: 164438 bases at least 430
Consensus quality: 164509 bases at least 420
Insert size: 161364; sum-of-contigs
Quality coverage: 12.88x in Q20 bases;sum-of-contigs

FEATURES
source 1. 161361
Location/Qualifiers
/organism="Homo sapiens"
/db_xref="taxon:9606"
/chromosome="3"
/map="3p"
/clone="RP11-127114"

BASE COUNT 46789 a 36594 c 35116 g 42862 t

ORIGIN

```

Query Match 1.6%: Score 24; DB 9; Length 161361;
Best Local Similarity 100.0%; Pred. No. 0.24;
Matches 24; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

1516 tctgcttcaaaaaaaaaa 1539
|||||
Db 115484 TCTGCTTCAAAAAAAAAA 115461

RESULT 13
AC026020/c 178622 bp DNA linear HTG 11-APR-2000
LOCUS Homo sapiens clone RP11-625C9, WORKING DRAFT SEQUENCE, 26 unordered
DEFINITION pieces.
ACCESSION AC026020
VERSION AC026020.2 GI:7534095
KEYWORDS HTG; HTGS_PHASE1; HTGS_DRAFT.
SOURCE human.
ORGANISM Homo sapiens

REFERENCE
AUTHORS
TITLE
JOURNAL
REFERENCE
AUTHORS

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.
1 (bases 1 to 178622)
Birren, B., Linton, L., Nusbaum, C. and Lander, E.
Homo sapiens, clone RP11-625C9
Unpublished
2 (bases 1 to 178622)
Birren, B., Linton, L., Nusbaum, C., Lander, E., Abraham, H., Allen, N., Anderson, S., Baldwin, J., Barna, N., Bastien, V., Bede, F., Boguslavsky, L., Boukhgalter, B., Brown, A., Burkett, G., Campoliano, A., Castle, A., Choepel, Y., Colangelo, M., Collins, S., Collymore, A., Cooke, P., Dearliano, K., Dewar, K., Diaz, J.S., Dodge, S., Domino, M., Doyle, M., Ferreira, P., Fitzhugh, W., Gage, D., Galagan, J., Gardyna, S., Ginde, S., Goyette, M., Graham, L., Grand-Pierre, N., Grant, C., Hagos, B., Heatford, A., Horton, L., Howland, J.C., Iliev, I., Johnson, R., Jones, C., Kann, L., Karatas, A., Klein, J., Lacroque, K., Lamazares, R., Landers, T., Lechoczky, J., Levine, R., Lieu, C., Liu, G., Locke, K., Macdonald, P., Marquis, N., McCarthy, M., McEwan, P., McGuirk, A., McKernan, K., McSheeters, R., Melidiri, J., Menus, L., Mihova, T., Miranda, C., Mienga, V., Morrow, J., Murphy, T., Naylor, J., Norman, C.H., O'Connor, T., O'Donnell, P., O'Neill, D., Oliver, T.M., Oliver, J., Peterson, K., Pierre, N., Pisanli, C., Pollara, V., Raymond, C., Riley, R., Rogov, P., Rothman, D., Roy, A., Santos, R., Schauer, S., Severy, P., Spencer, B., Stange-Thomann, N., Stojanovic, N., Subramanian, A., Talamas, J., Testaye, S., Theodore, J., Titrill, A., Travers, M., Trigilio, J., Vassiliev, H., Viel, R., Vo, A., Wilson, B., Wu, X., Wyman, D., Ye, W.J., Young, G., Zainoun, J., Zimmer, A. and Zody, M.

TITLE
JOURNAL
COMMENT

Submitted (19-MAR-2000) Whitehead Institute/MIT Center for Genome Research, 320 Charles Street, Cambridge, MA 02141, USA
On Apr 11, 2000 this sequence version replaced gi:7263239.
All repeats were identified using RepeatMasker:
Smit, A.F.A. & Green, P. (1996-1997)
http://ftp.genome.washington.edu/RM/RepeatMasker.html

Genome Center
Center: Whitehead Institute/ MIT Center for Genome Research
Center code: WIBR
Web site: http://www-seq.wi.mit.edu
Contact: sequence_submissions@genome.wi.mit.edu
Project Information
Center project name: L8574
Center clone name: 625_C_9

Summary Statistics
Sequencing vector: M13; M77815; 100% of reads
Chemistry: Dye-terminator Big Dye; 100% of reads
Assembly program: Phrap; version 0.960731
Consensus quality: 158834 bases at least Q40
Consensus quality: 169105 bases at least Q30
Consensus quality: 173272 bases at least Q20
Insert size: 189000; agarose-1p
Insert size: 176122; sum-of-coverage
Quality coverage: 3.4 in Q20 bases; agarose-1p
Quality coverage: 3.6 in Q20 bases; sum-of-coverage

NOTE: This is a 'working draft' sequence. It currently consists of 26 contigs. The true order of the pieces is not known and their order in this sequence record is arbitrary. Gaps between the contigs are represented as runs of N, but the exact sizes of the gaps are unknown. This record will be updated with the finished sequence as soon as it is available and the accession number will be preserved.

1 1098: contig of 1098 bp in length
* 1099 1198: gap of 100 bp
* 1199 2437: contig of 1239 bp in length
* 2438 2537: gap of 100 bp
* 2538 3838: contig of 1301 bp in length
* 3839 3938: gap of 100 bp
* 3939 5158: contig of 1220 bp in length
* 5159 5258: gap of 100 bp
* 5259 6847: contig of 1559 bp in length
* 6848 6947: gap of 100 bp

6948 8346: contig of 1399 bp in length
* 8347 8446: gap of 100 bp
* 8447 10954: contig of 2508 bp in length
* 10955 11054: gap of 100 bp
* 11055 13681: contig of 2627 bp in length
* 13682 13781: gap of 100 bp
* 13782 17774: contig of 3993 bp in length
* 17775 17874: gap of 100 bp
* 17875 21926: contig of 4052 bp in length
* 21927 22026: gap of 100 bp
* 22027 25801: contig of 3775 bp in length
* 25802 25901: gap of 100 bp
* 25902 30370: contig of 4469 bp in length
* 30371 30470: gap of 100 bp
* 30471 33259: contig of 2789 bp in length
* 33260 33359: gap of 100 bp
* 33360 36694: contig of 3335 bp in length
* 36695 36794: gap of 100 bp
* 36795 40643: contig of 3849 bp in length
* 40644 40743: gap of 100 bp
* 40744 46066: contig of 5323 bp in length
* 46067 46166: gap of 100 bp
* 46167 51967: contig of 5801 bp in length
* 51968 52067: gap of 100 bp
* 52068 59476: contig of 7409 bp in length
* 59477 59576: gap of 100 bp
* 59577 68880: contig of 9304 bp in length
* 68881 68980: gap of 100 bp
* 68981 78794: contig of 9814 bp in length
* 78795 78894: gap of 100 bp
* 78895 89939: contig of 11045 bp in length
* 89940 90039: gap of 100 bp
* 90040 101801: contig of 11762 bp in length
* 101802 101901: gap of 100 bp
* 101902 116968: contig of 15067 bp in length
* 116969 117068: gap of 100 bp
* 117069 132078: contig of 15010 bp in length
* 132079 132178: gap of 100 bp
* 132179 152928: contig of 20750 bp in length
* 152929 153028: gap of 100 bp
* 153029 178622: contig of 25594 bp in length.

FEATURES
source
1..178622
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="RP11-625C9"
/clone.lib="RP11 Human Male BAC"
1..1098
/note="assembly-fragment"
1199..2437
/note="assembly-fragment"
2538..3838
/note="assembly-fragment"
3939..5158
/note="assembly-fragment"
5259..6847
/note="assembly-fragment"
6948..8346
/note="assembly-fragment"
8447..10954
/note="assembly-fragment"
11055..13681
/note="assembly-fragment"
13782..17774
/note="assembly-fragment"
17875..21926
/note="assembly-fragment"
22027..25801
/note="assembly-fragment"
25902..30370
/note="assembly-fragment"
30471..33259
/note="assembly-fragment"

```

misc_feature      clone_end:T7
                  vector_side:right"
misc_feature      33360..36694
                  /note="assembly_fragment"
misc_feature      36795..40643
                  /note="assembly_fragment"
                  clone_end:SP6
                  vector_side:left"
misc_feature      40744..46066
                  /note="assembly_fragment"
misc_feature      46167..51967
                  /note="assembly_fragment"
misc_feature      52068..59476
                  /note="assembly_fragment"
misc_feature      59577..68880
                  /note="assembly_fragment"
misc_feature      68981..78794
                  /note="assembly_fragment"
misc_feature      78895..89939
                  /note="assembly_fragment"
misc_feature      90040..101801
                  /note="assembly_fragment"
misc_feature      101902..116968
                  /note="assembly_fragment"
misc_feature      117069..132078
                  /note="assembly_fragment"
misc_feature      132179..152928
                  /note="assembly_fragment"
misc_feature      153029..178622
                  /note="assembly_fragment"
BASE COUNT      50532 a 39135 c 38783 g 47653 t 2519 others
ORIGIN

```

```

Query Match      1.6%; Score 24; DB 2; Length 178622;
Best Local Similarity 100.0%; Pred. No. 0.24;
Matches 24; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

Oy 1516 tctgctcacaaaaa 1539
Db 42034 TCTGCTTACAAAAA 42011

```

```

RESULT 14
AL591488/c      191494 bp      DNA      linear      HTG 22-NOV-2001
LOCUS
DEFINITION      Mus musculus chromosome 2 clone RP23-36P22, *** SEQUENCING IN
ACCESSION      AL591488
VERSION      AL591488.7 GI:17065727
KEYWORDS      HTG; HTGS_PHASE1; HTGS_ACTIVEFIN; HTGS_DRAFT; HTGS_FULLTOP.
SOURCE      house mouse.
ORGANISM      Mus musculus
                  Eukaryota; Chordata; Craniata; Vertebrata; Euteleostomi;
                  Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE
AUTHORS      Pearce,A
TITLE      Direct Submission
JOURNAL      Submitted (21-NOV-2001) Wellcome Trust Sanger Institute, Hinxton,
                  Cambridgeshire, CB10 1SA, UK. E-mail enquiries:
                  humquerry@sanger.ac.uk Clone requests: clonerequest@sanger.ac.uk
                  On Nov 25, 2001 this sequence version replaced gi:17043820.
COMMENT
                  ----- Genome Center
                  Center: Wellcome Trust Sanger Institute
                  Center code: SC
                  Web site: http://www.sanger.ac.uk
                  Contact: humquerry@sanger.ac.uk
                  ----- Project Information
                  Center project name: BM36P22
                  ----- Summary Statistics
                  Assembly program: XGAP4; version 4.5
                  Sequencing vector: plasmid; L08752; 100% of reads
                  Chemistry: Dye-terminator Big Dye; 99% of reads

```

```

Chemistry: Dye-primer Big Dye; 0% of reads
Consensus quality: 191473 bases at least Q40
Consensus quality: 191494 bases at least Q30
Consensus quality: 191494 bases at least Q20
Insert size: 191494; sum-of-contigs
Insert size: 168060; 8.7% error; agarose-fp
Quality coverage: 16.18x in Q20 bases; sum-of-contigs Quality
coverage: 18.44x in Q20 bases; agarose-fp
* NOTE: This is a 'working draft' sequence.
* This sequence will be replaced
* by the finished sequence as soon as it is available and
* the accession number will be preserved.
FEATURES
source
1..191494
/organism="Mus musculus"
/db_xref="taxon:10090"
/chromosome="2"
/clone="RP23-36P22"
/clone.lib="RPC1-23"
misc_feature 1..191494
/note="assembly_fragment:03311"
BASE COUNT      49688 a 46072 c 46038 g 49696 t
ORIGIN

```

```

Query Match      1.6%; Score 24; DB 2; Length 191494;
Best Local Similarity 100.0%; Pred. No. 0.24;
Matches 24; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

Oy 1516 tctgctcacaaaaa 1539
Db 66072 TCTGCTTACAAAAA 66049

```

```

RESULT 15
AL591607/c      239135 bp      DNA      linear      HTG 13-NOV-2001
LOCUS
DEFINITION      Mus musculus chromosome 2 clone RP23-215C14, *** SEQUENCING IN
ACCESSION      AL591607
VERSION      AL591607.5 GI:16944832
KEYWORDS      HTG; HTGS_PHASE1; HTGS_DRAFT; HTGS_FULLTOP.
SOURCE      house mouse.
ORGANISM      Mus musculus
                  Eukaryota; Chordata; Craniata; Vertebrata; Euteleostomi;
                  Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE
AUTHORS      Sims,S.
TITLE      Direct Submission
JOURNAL      Submitted (10-NOV-2001) Wellcome Trust Sanger Institute, Hinxton,
                  Cambridgeshire, CB10 1SA, UK. E-mail enquiries:
                  humquerry@sanger.ac.uk Clone requests: clonerequest@sanger.ac.uk
                  On Nov 15, 2001 this sequence version replaced gi:14330183.
COMMENT
                  ----- Genome Center
                  Center: Wellcome Trust Sanger Institute
                  Center code: SC
                  Web site: http://www.sanger.ac.uk
                  Contact: humquerry@sanger.ac.uk
                  ----- Project Information
                  Center project name: BM215C14
                  ----- Summary Statistics
                  Assembly program: XGAP4; version 4.5
                  Sequencing vector: plasmid; L08752; 100% of reads
                  Chemistry: Dye-terminator Big Dye; 100% of reads
                  Consensus quality: 236982 bases at least Q40
                  Consensus quality: 237392 bases at least Q30
                  Consensus quality: 237788 bases at least Q20
                  Insert size: 238235; sum-of-contigs
                  Insert size: 240508; 3.8% error; agarose-fp
                  Quality coverage: 15.35x in Q20 bases; sum-of-contigs Quality
                  coverage: 15.29x in Q20 bases; agarose-fp

```


GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: July 18, 2002, 10:19:37 ; Search time 1762.48 Seconds
(without alignments)
11785.560 Million cell updates/sec

Title: US-09-732-680A-1
Perfect score: 1539
Sequence: 1 attcgacagcagggatcac.....cttcaaaaaaaaaaaaaa 1539

Scoring table: OLIGO_NUC
Gapop 60.0 , Gapext 60.0

Searched: 13736207 segs, 674847542 residues

rd size : 0

Total number of hits satisfying chosen parameters: 27472414

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Listing first 1000 summaries

Database :

EST:*
1: em_estba:*
2: em_esthum:*
3: em_estin:*
4: em_estnu:*
5: em_estov:*
6: em_estpl:*
7: em_estro:*
8: em_hlc:*
9: gb_estl:*
10: gb_est2:*
11: gb_hlc:*
12: gb_gss:*
13: em_gss_hum:*
14: em_gss_inu:*
15: em_gss_pln:*
16: em_gss_vrt:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	288	18.7	448	9	AI542842 SD09326.5
2	162	10.5	890	12	CNS0137Y
3	57	3.7	361	9	AM944588 SD09326.3
4	24	1.6	303	10	BI285394
5	24	1.6	445	9	AI801505
6	24	1.6	475	10	BI399302
7	24	1.6	520	10	BI285223
8	24	1.6	548	12	AZ588949
9	24	1.6	1560	11	AK017944
10	23	1.5	198	10	BG939598
11	23	1.5	347	9	AI535508
12	23	1.5	351	9	AI574766
13	23	1.5	359	10	BF703043
14	23	1.5	360	10	D68664
15	23	1.5	360	10	D69343
16	23	1.5	372	9	AI853953
17	23	1.5	402	9	AI955429

C 18	23	1.5	423	9	AM520985
C 19	23	1.5	447	12	BH593220
C 20	23	1.5	474	9	AM276309
C 21	23	1.5	486	10	BI896248
C 22	23	1.5	493	10	BF401313
C 23	23	1.5	512	10	BE655378
C 24	23	1.5	525	9	AA899747
C 25	23	1.5	572	10	BM428786
C 26	23	1.5	593	10	BI119595
C 27	23	1.5	603	10	BM428787
C 28	23	1.5	621	10	BG436490
C 29	23	1.5	623	9	AM015683
C 30	23	1.5	696	10	BI668117
C 31	22	1.4	75	9	AA615120
C 32	22	1.4	103	10	BG732100
C 33	22	1.4	126	10	BG375749
C 34	22	1.4	139	10	BF057920
C 35	22	1.4	157	10	BM033721
C 36	22	1.4	160	9	AI349175
C 37	22	1.4	174	9	AI251258
C 38	22	1.4	206	10	BM148619
C 39	22	1.4	224	9	AA964311
C 40	22	1.4	254	9	AM822353
C 41	22	1.4	275	10	BM379065
C 42	22	1.4	275	10	BM379104
C 43	22	1.4	282	10	BF511708
C 44	22	1.4	312	9	AJ285166
C 45	22	1.4	313	9	AI454684
C 46	22	1.4	313	10	BF270202
C 47	22	1.4	315	10	BE802075
C 48	22	1.4	326	10	BF557263
C 49	22	1.4	327	10	BM380422
C 50	22	1.4	353	9	AJ285397
C 51	22	1.4	360	9	AM298450
C 52	22	1.4	364	9	AM045460
C 53	22	1.4	388	12	AZ321692
C 54	22	1.4	394	9	AM298730
C 55	22	1.4	411	9	AM532526
C 56	22	1.4	413	9	BE103145
C 57	22	1.4	416	10	BE992989
C 58	22	1.4	419	9	AM738838
C 59	22	1.4	434	10	BE990134
C 60	22	1.4	436	9	AA052898
C 61	22	1.4	439	9	AM520090
C 62	22	1.4	444	10	BI302642
C 63	22	1.4	466	9	AI840647
C 64	22	1.4	475	9	BE110024
C 65	22	1.4	490	9	AI848123
C 66	22	1.4	511	9	AA964637
C 67	22	1.4	532	12	AO589092
C 68	22	1.4	535	10	BE426242
C 69	22	1.4	548	10	BI133496
C 70	22	1.4	554	12	AZ825313
C 71	22	1.4	557	10	BG841202
C 72	22	1.4	568	9	AUI65548
C 73	22	1.4	581	10	BM382438
C 74	22	1.4	611	12	AZ290954
C 75	22	1.4	632	9	AM327740
C 76	22	1.4	661	10	BM380561
C 77	22	1.4	676	10	BM379176
C 78	22	1.4	680	9	AU067768
C 79	22	1.4	681	10	BF635116
C 80	22	1.4	857	10	BE676634
C 81	21	1.4	73	10	BG621583
C 82	21	1.4	116	10	BG362096
C 83	21	1.4	161	10	BG272467
C 84	21	1.4	169	12	B80272
C 85	21	1.4	176	9	BE043547
C 86	21	1.4	187	9	AI349627
C 87	21	1.4	188	9	AI250383
C 88	21	1.4	190	9	AI349135
C 89	21	1.4	207	9	AI581229
C 90	21	1.4	210	9	BB609102

AM520985	UI-R-BOO-
BH593220	BOHRM35TF
AM276309	xt10a10.x
BI896248	td54n06.x
BF401313	UI-R-CAO-
BE655378	UI-M-BHO-
AA899747	UI-R-EO-d
BM428786	952023D10
BI119595	952023D10
BM428787	952023D10
BG436490	602509162
AM015683	UI-H-BIOP
BI668117	603295951
AA615120	vn02b09.r
BG732100	ps20c04.y
BG375749	UI-R-CSO-
BF057920	7176h08.y
BM033721	kh69g01.y
AI349175	ta73d10.x
AI251258	qv36g03.x
BM148619	TCNAP2D63
AA964311	UI-R-CU-g
AM822353	uq18f09.x
BM379065	MEST573-C
BM379104	MEST573-C
BF511708	UI-H-BI4-
AJ285166	4A3B-AAG-
AI454684	UI-R-BT0-
BF270202	GA_EB000
BE802075	sr23g07.y
BF557263	UI-R-C0-g
BM380422	MEST519-F
AJ285397	4A3B-AAT-
AM298450	UI-H-BW0-
AM045460	UI-M-BH1-
AZ321692	1M0042C24
AM298730	UI-H-BW0-
AM532526	UI-R-BSO-
BE103145	UI-R-BT1-
BE992989	UI-M-BZ1-
AM738838	gpb03b01.y
BE990134	UI-M-BZ1-
AA052898	z171g03.s
AM520090	660056H04
BI302642	UI-R-DMO-
AI840647	UI-M-AM0-
BE110024	UI-R-CAO-
AI848123	UI-M-AP1-
AA964637	UI-R-C0-h
AO589092	CITBT-EI-
BE426242	WHE0329_A
BI133496	UI-M-BH3-
AZ825313	2M0100L24
BG841202	MEST18-C0
AUI65548	AUI65548
BM382438	MEST549-F
AZ290954	RPCI-23-1
AM327740	d10f110.x
BM380561	MEST521-F
BM379176	MEST500-G
AU067768	AU067768
BF635116	hR079h02D
BE676634	602086484
BG621583	602617180
BG362096	gbd5b06.y
BG272467	nah33b12.
BE043547	hR440g07.y
AI349627	ta74h09.x
AI250383	gx23f03.x
AI349135	ta51h09.x
AI581229	td70h08.x
BB609102	BB609102

c 91	21	1.4	227	10	BM390422	BM390422	UI-R-CNI-	164	20	1.3	102	10	BI682371	BI682371	463436	MA
c 92	21	1.4	231	10	BI341453	BI341453	368885	MA	c 165	20	1.3	116	10	BE634750	BE634750	uv86e12.x
c 93	21	1.4	240	9	AM314351	AM314351	10381	MAR	c 166	20	1.3	118	9	AI463121	AI463121	uc44c03.x
c 94	21	1.4	254	9	AV735874	AV735874	AV735874		c 167	20	1.3	119	10	BI615187	BI615187	na137b07.
c 95	21	1.4	257	9	AM071389	AM071389	xa64g01.x		c 168	20	1.3	119	10	BE511759	BE511759	UI-H-BT4-
c 96	21	1.4	301	9	AI089807	AI089807	qa11b08.x		c 169	20	1.3	121	10	BI293609	BI293609	UI-R-DK0-
c 97	21	1.4	301	9	AM294743	AM294743	UI-H-BW0-		c 170	20	1.3	125	10	BI6187569	BI6187569	RS76698.A
c 98	21	1.4	303	9	AI562632	AI562632	TENS2618		c 171	20	1.3	126	10	BE955288	BE955288	UI-M-BH4-
c 99	21	1.4	305	9	AI159836	AI159836	qB51a08.x		c 172	20	1.3	127	10	BE583559	BE583559	3-7D-HA.P
c 100	21	1.4	308	10	BE377237	BE377237	UI-R-CU0-		c 173	20	1.3	133	9	AM021567	AM021567	dE25d03.Y
c 101	21	1.4	314	9	BE110420	BE110420	UI-R-BJ1-		c 174	20	1.3	137	9	BE116928	BE116928	UI-R-B51-
c 102	21	1.4	320	10	Z36865	Z36865	HHR663	Att	c 175	20	1.3	138	9	AM139339	AM139339	UI-H-BT1-
c 103	21	1.4	325	12	AI263539	AI263539	ZM0232C16		c 176	20	1.3	156	10	BE990522	BE990522	UI-M-B21-
c 104	21	1.4	327	10	BE509458	BE509458	UI-H-B14-		c 177	20	1.3	160	10	BG940656	BG940656	ax07a01.Y
c 105	21	1.4	331	9	AI535265	AI535265	UI-R-C3-S		c 178	20	1.3	162	9	AI792550	AI792550	q17g02.Y
c 106	21	1.4	362	9	AA836519	AA836519	od22d03.S		c 179	20	1.3	163	9	AM397667	AM397667	sg82609.Y
c 107	21	1.4	368	10	BF459090	BF459090	UI-M-B21-		c 180	20	1.3	164	9	AI136641	AI136641	UI-R-C2P-
c 108	21	1.4	369	9	AI145074	AI145074	UI-R-BT0-		c 181	20	1.3	169	9	AI305981	AI305981	qW48g11.x
c 109	21	1.4	371	9	AI285066	AI285066	AA3B-AAF-		c 182	20	1.3	171	9	BI474953	BI474953	IP35c01.x
c 110	21	1.4	381	10	BE557416	BE557416	W95h06.x		c 183	20	1.3	171	9	AA925693	AA925693	UI-R-A1-e
c 111	21	1.4	393	10	BE557416	BE557416	UI-R-B12.Y		c 184	20	1.3	171	9	AI436532	AI436532	th52c01.x
c 112	21	1.4	394	10	BF408009	BF408009	UI-R-BJ2-		c 185	20	1.3	171	10	BM385896	BM385896	UI-R-DM1-
c 113	21	1.4	404	10	BE333206	BE333206	uT01g01.Y		c 186	20	1.3	174	9	AI311029	AI311029	ta45e04.x
c 114	21	1.4	417	9	AM047241	AM047241	UI-M-BH1-		c 187	20	1.3	175	9	AI399123	AI399123	NCM13A12T
c 115	21	1.4	424	9	AI699030	AI699030	TX74D08.x		c 188	20	1.3	175	9	AI613361	AI613361	LT81c06.x
c 116	21	1.4	426	10	BI503994	BI503994	BI170024B		c 189	20	1.3	176	9	AI251527	AI251527	qV28a08.x
c 117	21	1.4	439	9	AI070882	AI070882	UI-R-C2-M		c 190	20	1.3	178	9	AI254848	AI254848	qV36a07.x
c 118	21	1.4	441	9	AI201294	AI201294	qf65h12.x		c 191	20	1.3	178	9	AI311432	AI311432	ta45c01.x
c 119	21	1.4	454	10	BM388354	BM388354	UI-R-CN1-		c 192	20	1.3	179	9	AI254370	AI254370	qV33b03.x
c 120	21	1.4	456	9	AA754029	AA754029	97BS0497		c 193	20	1.3	179	10	BE981135	BE981135	UI-M-CG0-
c 121	21	1.4	464	9	BE026336	BE026336	dB37d04.Y		c 194	20	1.3	182	9	BE042120	BE042120	h15a04.Y
c 122	21	1.4	474	9	AI365649	AI365649	MCBA01F07		c 195	20	1.3	182	10	BF986605	BF986605	CM0-GN10
c 123	21	1.4	474	9	AM067248	AM067248	683019F10		c 196	20	1.3	182	10	BF994984	BF994984	CM0-GN10
c 124	21	1.4	480	10	BI295392	BI295392	UI-R-DK0-		c 197	20	1.3	183	9	AI366174	AI366174	MEBA05D08
c 125	21	1.4	487	9	AI097287	AI097287	qB88f12.x		c 198	20	1.3	184	9	AI313345	AI313345	ta77h04.x
c 126	21	1.4	492	10	BE409253	BE409253	UI-R-BT1-		c 199	20	1.3	184	9	AV739755	AV739755	AV739755
c 127	21	1.4	498	10	BM385795	BM385795	UI-R-CN1-		c 200	20	1.3	184	10	BF250279	BF250279	pa97b06.Y
c 128	21	1.4	501	9	AM061064	AM061064	UI-M-BH1-		c 201	20	1.3	185	9	AI308397	AI308397	tb43d09.x
c 129	21	1.4	503	9	AM951298	AM951298	EST363368		c 202	20	1.3	185	9	AA145807	AA145807	mr32d10.x
c 130	21	1.4	524	9	AL384959	AL384959	MEBC25E07		c 203	20	1.3	185	9	AV735467	AV735467	AV735467
c 131	21	1.4	526	10	BI796115	BI796115	H038F05.E		c 204	20	1.3	186	9	AM204754	AM204754	UI-H-BU1-
c 132	21	1.4	528	9	AL368733	AL368733	MEBA26D02		c 205	20	1.3	187	9	AM297442	AM297442	UI-H-BW0-
c 133	21	1.4	568	10	BI274264	BI274264	UI-R-CW0-		c 206	20	1.3	188	9	AM654270	AM654270	103603.MA
c 134	21	1.4	585	10	BE268694	BE268694	1000208D0		c 207	20	1.3	189	9	AM302047	AM302047	xt52f08.x
c 135	21	1.4	586	10	BM379871	BM379871	ME5711-G		c 208	20	1.3	190	9	BE114657	BE114657	UI-R-BJ1-
c 136	21	1.4	600	9	AA898269	AA898269	NCC2A12T7		c 209	20	1.3	191	9	AI029192	AI029192	UI-R-C0-1
c 137	21	1.4	602	9	AI722145	AI722145	f319g08.Y		c 210	20	1.3	193	9	AI312498	AI312498	ta91g10.x
c 138	21	1.4	608	10	BI290555	BI290555	UI-R-DK0-		c 211	20	1.3	193	10	BI945735	BI945735	sc66e07.Y
c 139	21	1.4	626	9	AV694025	AV694025	AV654055		c 212	20	1.3	195	9	AM120595	AM120595	UI-M-BH2.
c 140	21	1.4	634	10	BE575184	BE575184	602597719		c 213	20	1.3	198	10	BI321192	BI321192	sa148f05.
c 141	21	1.4	637	10	BE503105	BE503105	602551089		c 214	20	1.3	201	9	AI836404	AI836404	UI-M-AP0-
c 142	21	1.4	640	10	BI798154	BI798154	H099B04.E		c 215	20	1.3	203	10	BE706513	BE706513	UI-R-BJ1-
c 143	21	1.4	644	9	AV685350	AV685350	AV685350		c 216	20	1.3	204	10	BI280016	BI280016	UI-R-DD0-
c 144	21	1.4	667	10	BF684272	BF684272	602141190		c 217	20	1.3	205	9	AI192763	AI192763	qe62g12.x
c 145	21	1.4	689	9	AU092678	AU092678	AU092678		c 218	20	1.3	206	9	AI611528	AI611528	LT65f01.x
c 146	21	1.4	699	10	BE731652	BE731652	601567032		c 219	20	1.3	207	9	AM490702	AM490702	UI-M-BH3-
c 147	21	1.4	700	9	AL501893	AL501893	AL501893		c 220	20	1.3	207	10	BF019259	BF019259	fy03a11.Y
c 148	21	1.4	702	10	BE254472	BE254472	601111562		c 221	20	1.3	211	10	BM026482	BM026482	uF05a09.x
c 149	21	1.4	711	10	BE640956	BE640956	MEST13-H0		c 222	20	1.3	215	9	AL366090	AL366090	h017e08.x
c 150	21	1.4	715	10	BE267019	BE267019	601192680		c 223	20	1.3	215	9	AM779032	AM779032	h017e08.x
c 151	21	1.4	723	10	BE260856	BE260856	601151266		c 224	20	1.3	218	10	BE162062	BE162062	dc45c02.Y
c 152	21	1.4	757	10	BI853586	BI853586	601380709		c 225	20	1.3	219	9	AI932631	AI932631	wo05e05.x
c 153	21	1.4	760	10	BE514673	BE514673	60131737		c 226	20	1.3	219	9	AI986680	AI986680	rs20b11.Y
c 154	21	1.4	777	12	BH510568	BH510568	BOHR05TE		c 227	20	1.3	221	9	BE049186	BE049186	fa73e08.x
c 155	21	1.4	797	10	BE334782	BE334782	602460979		c 228	20	1.3	221	9	AA494912	AA494912	fa12g04.s
c 156	21	1.4	922	12	CMS05T00	CMS05T00	Tetraodon		c 229	20	1.3	225	9	AI252894	AI252894	qv31c12.x
c 157	21	1.4	937	12	BM015808	BM015808	603642074		c 230	20	1.3	225	10	BE953771	BE953771	UI-M-CC1-
c 158	21	1.4	1272	12	AG082552	AG082552	Pan t10g1		c 231	20	1.3	226	9	AI251068	AI251068	qx76a08.x
c 159	21	1.4	2747	11	BC013664	BC013664	Mus muscu		c 232	20	1.3	226	9	AM078837	AM078837	xb17h09.x
c 160	20	1.3	48	10	BM144540	BM144540	TCAP1D11		c 233	20	1.3	227	10	BE139372	BE139372	xr69f07.x
c 161	20	1.3	98	9	AI254470	AI254470	qV47h04.x		c 234	20	1.3	227	10	BF992655	BF992655	UI-R-CM0-
c 162	20	1.3	98	9	AM311302	AM311302	sg35b11.Y		c 235	20	1.3	228	9	AI986114	AI986114	wR0e03.x
c 163	20	1.3	102	9	AM457474	AM457474	UI-M-BH3-		c 236	20	1.3	232	10	BE684622	BE684622	186255.MA

C 237	20	1.3	234	9	AA595937	AA595937	nm6b05..s	310	20	1.3	300	10	BE584667	BE584667	5-10b-ZO
C 238	20	1.3	235	9	AU053650	AU053650	AU053650	311	20	1.3	302	9	AV737425	AV737425	AV737425
C 239	20	1.3	236	9	AA433561	AA433561	UT-R-B10P	312	20	1.3	303	9	AI452006	AI452006	AI452006
C 240	20	1.3	237	9	AM304297	AM304297	xy58d07.x	313	20	1.3	303	9	AM904833	AM904833	RCS-NM106
C 241	20	1.3	240	9	AI361630	AI361630	qy86c04.x	314	20	1.3	304	9	AA714031	AA714031	nx81f09..s
C 242	20	1.3	241	10	BG362300	BG362300	gy70d11.y	315	20	1.3	305	9	AI850824	AI850824	UT-M-BG1-
C 243	20	1.3	242	10	BM087542	BM087542	500327.MA	316	20	1.3	305	10	BG880093	BG880093	1b74d08..x
C 244	20	1.3	243	9	AI839779	AI839779	UT-M-ANO-	317	20	1.3	306	9	AA586909	AA586909	nm68e05..s
C 245	20	1.3	245	9	AA260643	AA260643	va12f11.r	318	20	1.3	306	10	BF406038	BF406038	UT-R-CAI-
C 246	20	1.3	245	12	AQ040790	AQ040790	CIT-HSP-2	319	20	1.3	308	9	AA925564	AA925564	UT-R-AI-e
C 247	20	1.3	247	9	AV741838	AV741838	BF469006	320	20	1.3	310	9	AI073243	AI073243	UT-R-YO-1
C 248	20	1.3	247	10	BF469006	BF469006	UT-M-BH3-	321	20	1.3	310	9	BB099461	BB099461	BB099461
C 249	20	1.3	248	9	AU101315	AU101315	AU101315	322	20	1.3	311	9	AI059421	AI059421	UT-R-C1-K
C 250	20	1.3	248	10	BI535631	BI535631	399268.MA	322	20	1.3	311	9	BE115146	BE115146	UT-M-BS1-
C 251	20	1.3	248	10	C22109	C22109	M4YA	323	20	1.3	312	10	BI535572	BI535572	399182.MA
C 252	20	1.3	249	10	BE530645	BE530645	M781235TM	325	20	1.3	313	10	BE989971	BE989971	UT-M-BZ1-
C 253	20	1.3	250	9	AI853431	AI853431	UT-M-BH0-	326	20	1.3	314	9	AA555317	AA555317	AV555317
C 254	20	1.3	250	10	BE949196	BE949196	UT-M-BH3-	327	20	1.3	315	9	AA406310	AA406310	zv24c11..s
C 255	20	1.3	252	9	AV373870	AV373870	AV373870	328	20	1.3	315	10	BE985968	BE985968	UT-M-CGDP
C 256	20	1.3	253	9	AI763501	AI763501	UT-R-YO-a	329	20	1.3	315	10	BE986834	BE986834	UT-M-CGDP
C 257	20	1.3	253	9	AI856083	AI856083	sg31d07.x	330	20	1.3	315	10	BE989720	BE989720	UT-M-BZ1-
C 258	20	1.3	253	9	AM294507	AM294507	UT-H-BI2-	331	20	1.3	317	10	BM153387	BM153387	TCBAP2D11
C 259	20	1.3	253	10	BF392335	BF392335	UT-R-CAO-	332	20	1.3	318	10	BG385066	BG385066	306656.MA
C 260	20	1.3	254	9	BG018500	BG018500	daa46d06..	333	20	1.3	321	9	AI098578	AI098578	ue29g11..x
C 261	20	1.3	255	9	AI310919	AI310919	ta79e10..x	334	20	1.3	321	9	AA252799	AA252799	UT-R-B10-
C 262	20	1.3	255	9	AI312243	AI312243	rv20d08..s	335	20	1.3	323	9	AA116810	AA116810	MM28e03..r
C 263	20	1.3	257	9	AA478793	AA478793	TCGAP1D83	336	20	1.3	323	9	AI714209	AI714209	UT-R-AE1-
C 264	20	1.3	260	10	BM193791	BM193791	32DS.AmerB	337	20	1.3	325	9	BE107137	BE107137	UT-R-BS1-
C 265	20	1.3	260	10	BF049568	BF049568	qk76g10..x	338	20	1.3	326	9	AM528617	AM528617	UT-R-B01-
C 266	20	1.3	262	9	AI249642	AI249642	UT-R-CAI-	339	20	1.3	327	9	AI117576	AI117576	u143609..x
C 267	20	1.3	263	10	BF388379	BF388379	UT-R-CAI-	340	20	1.3	327	9	AW703191	AW703191	T9ESTx284
C 268	20	1.3	265	9	AI311422	AI311422	ta45a11..x	341	20	1.3	327	10	BI276342	BI276342	UT-R-CW0-
C 269	20	1.3	266	9	AV742051	AV742051	R64886..13390	342	20	1.3	327	10	BE987503	BE987503	UT-M-CGDP
C 270	20	1.3	266	10	R64886	R64886	13390.lambd	343	20	1.3	329	9	AW345723	AW345723	15592.MAR
C 271	20	1.3	267	9	AV739381	AV739381	AV739381	344	20	1.3	329	10	D22540	D22540	RICCO233A.R
C 272	20	1.3	267	10	BF049570	BF049570	98/700.Am	345	20	1.3	330	9	AA829834	AA829834	0d40c03..s
C 273	20	1.3	268	9	AI583821	AI583821	te73b09..x	346	20	1.3	330	9	AI547986	AI547986	UT-R-C3-L
C 274	20	1.3	269	9	AI044518	AI044518	UT-R-C1-K	347	20	1.3	330	9	AM489749	AM489749	UT-M-BH3-
C 275	20	1.3	269	10	BM465647	BM465647	AGENCOUNT	348	20	1.3	330	10	R39624	R39624	yh6b01.r1
C 276	20	1.3	270	9	AM323007	AM323007	uo56b11..x	349	20	1.3	331	9	AM449165	AM449165	UT-H-BI3-
C 277	20	1.3	270	10	BF415779	BF415779	UT-R-CAO-	350	20	1.3	331	10	BI694013	BI694013	UT-R-GO-u
C 278	20	1.3	272	9	AI340518	AI340518	tb30c01..x	351	20	1.3	332	9	AI575131	AI575131	UT-R-GO-u
C 279	20	1.3	272	9	AV740587	AV740587	BB020648	352	20	1.3	332	9	AA297065	AA297065	UT-M-BW0-
C 280	20	1.3	272	9	BB020648	BB020648	BB058442	353	20	1.3	335	10	BG944551	BG944551	ax51g11..x
C 281	20	1.3	272	9	BB058442	BB058442	MT-P-CPI-	354	20	1.3	336	9	AI851898	AI851898	UT-M-BH0-
C 282	20	1.3	272	10	BI403748	BI403748	AV116578	355	20	1.3	336	9	AU069910	AU069910	AU069910
C 283	20	1.3	274	9	AV116578	AV116578	BI804100	356	20	1.3	337	9	AV742675	AV742675	AV742675
C 284	20	1.3	274	10	BI804100	BI804100	H127C01.E	357	20	1.3	337	9	AI028959	AI028959	UT-R-CO-1
C 285	20	1.3	275	9	AI711026	AI711026	UT-R-AE1-	358	20	1.3	338	9	AI703814	AI703814	UT-R-AC1-
C 286	20	1.3	275	10	BF793539	BF793539	602255056	359	20	1.3	338	9	AM434201	AM434201	UT-R-B10P
C 287	20	1.3	277	9	AL659459	AL659459	UT-R-CS0S	360	20	1.3	339	12	A2552303	A2552303	RPCT-23-2
C 288	20	1.3	279	10	BI287759	BI287759	UT-R-CS0S	361	20	1.3	340	9	AI844015	AI844015	UT-M-AL1-
C 289	20	1.3	280	9	AI030319	AI030319	UT-R-CO-1	362	20	1.3	341	9	AI842121	AI842121	UT-M-AN1-
C 290	20	1.3	280	9	AI311624	AI311624	ta43g06..x	363	20	1.3	341	9	AM099119	AM099119	sg34f12..y
C 291	20	1.3	282	10	BI007372	BI007372	MRI-RT007	364	20	1.3	341	10	BG731992	BG731992	ps21b04..y
C 292	20	1.3	284	10	BI688659	BI688659	603311280	365	20	1.3	341	10	H10028	H10028	ymb2h11..s1
C 293	20	1.3	285	9	AI463565	AI463565	v6g3c09..x	366	20	1.3	343	9	AI553370	AI553370	UT-R-C3-s
C 294	20	1.3	285	10	BI430970	BI430970	9490635E08	367	20	1.3	343	9	AM158360	AM158360	z4d0002..x
C 295	20	1.3	285	10	C22976	C22976	D1cct	368	20	1.3	344	9	AA028165	AA028165	ze75e10..r
C 296	20	1.3	288	10	BM279585	BM279585	71/700.Am	369	20	1.3	344	9	AI056027	AI056027	ox46a01..s
C 297	20	1.3	289	9	AI547649	AI547649	UT-R-C3-s	370	20	1.3	344	9	AI705027	AI705027	UT-R-AB1-
C 298	20	1.3	289	9	AI841050	AI841050	UT-M-AMO-	371	20	1.3	344	9	AM910018	AM910018	ur78b12..y
C 299	20	1.3	290	9	AA565802	AA565802	nk26c07..s	372	20	1.3	344	10	BM383300	BM383300	UT-R-DS0-
C 300	20	1.3	291	9	AI611260	AI611260	tu14h11..x	373	20	1.3	345	9	AJ280951	AJ280951	4A3A-AAZ-
C 301	20	1.3	291	9	BE023949	BE023949	sm94f01..y	374	20	1.3	346	10	BG361998	BG361998	gb50a06..y
C 302	20	1.3	292	9	AA955117	AA955117	UT-R-A1-e	375	20	1.3	347	9	AI637031	AI637031	496003D11
C 303	20	1.3	292	9	BB070625	BB070625	BE115220	376	20	1.3	348	9	AA960497	AA960497	wv63d10..s
C 304	20	1.3	293	9	BE115220	BE115220	UT-R-BJ1-	377	20	1.3	348	9	AI686698	AI686698	tuc3h11..x
C 305	20	1.3	295	9	AV743970	AV743970	UT-M-CGDP	378	20	1.3	349	9	AI731642	AI731642	BMtGH103
C 306	20	1.3	295	10	BF467202	BF467202	UT-M-CGDP	379	20	1.3	349	10	BF704512	BF704512	MT-P-H4-a
C 307	20	1.3	297	10	BG372285	BG372285	UT-R-CV0-	380	20	1.3	350	10	BG734845	BG734845	CC-esf1cL
C 308	20	1.3	298	9	AA476285	AA476285	w295b09..s	381	20	1.3	351	9	AI840625	AI840625	UT-M-AMO-
C 309	20	1.3	298	10	BF074600	BF074600	222056.MA	382	20	1.3	351	9	AM597704	AM597704	sj49g07..y

C 383	20	1.3	355	10	B1298246	B1298246	UI-R-CV2-
C 384	20	1.3	355	10	BF397455	BF397455	UI-R-B52-
C 385	20	1.3	355	10	BF513273	BF513273	UI-H-BM1-
C 386	20	1.3	357	9	AM121811	AM121811	UI-M-BH2-
C 387	20	1.3	358	9	AI454747	AI454747	UI-R-C2P-
C 388	20	1.3	358	9	A1713509	A1713509	UI-R-AA1-
C 389	20	1.3	358	9	A1716632	A1716632	UI-R-Y0-a
C 390	20	1.3	358	9	AV738725	AV738725	AV738725
C 391	20	1.3	358	10	B1942310	B1942310	sh01f04.y
C 392	20	1.3	358	10	BF388140	BF388140	UI-R-CM1-
C 393	20	1.3	359	9	AI555190	AI555190	UI-R-C2P-
C 394	20	1.3	360	10	BG553087	BG553087	dab83602.
C 395	20	1.3	361	9	AV741090	AV741090	AV741090
C 396	20	1.3	361	9	AV741092	AV741092	AV741092
C 397	20	1.3	361	10	B1118928	B1118928	AR20H2.Po
C 398	20	1.3	362	9	AV739589	AV739589	AV739589
C 399	20	1.3	362	9	AM015593	AM015593	UI-H-B10-
C 400	20	1.3	362	10	BG671410	BG671410	DRNB0G06
C 401	20	1.3	362	10	B1404266	B1404266	MI-P-CP1-
C 402	20	1.3	363	9	AM046223	AM046223	UI-M-BH1-
C 403	20	1.3	363	9	AM060472	AM060472	UI-M-BH1-
C 405	20	1.3	364	9	AM523003	AM523003	UI-R-B00-
C 406	20	1.3	364	10	BF418467	BF418467	UI-R-B12-
C 407	20	1.3	365	10	B1281516	B1281516	UI-R-BD0-
C 408	20	1.3	366	9	AM028405	AM028405	w88901.x
C 409	20	1.3	366	10	BE990575	BE990575	UI-M-B21-
C 410	20	1.3	366	10	BF512429	BF512429	UI-H-B13-
C 411	20	1.3	368	10	B1286153	B1286153	UI-R-DE0-
C 412	20	1.3	369	9	AA899709	AA899709	UI-R-E0-d
C 413	20	1.3	369	9	AI044563	AI044563	UI-R-Cl-1
C 414	20	1.3	370	9	AL379190	AL379190	MEB843G06
C 415	20	1.3	370	10	BG362520	BG362520	gb73901.y
C 416	20	1.3	370	10	B7059835	B7059835	B7059835
C 417	20	1.3	370	10	BE994680	BE994680	UI-M-B21-
C 418	20	1.3	370	10	BR455356	BR455356	UI-M-B21-
C 419	20	1.3	371	9	AM122297	AM122297	UI-M-BH2-
C 420	20	1.3	372	9	AI463568	AI463568	v63d10.x
C 421	20	1.3	372	10	BM321105	BM321105	rockefell
C 422	20	1.3	373	9	AL386807	AL386807	MEB37A01
C 423	20	1.3	373	9	AA452663	AA452663	UI-H-B13-
C 424	20	1.3	374	10	BR408343	BR408343	UI-R-B12-
C 425	20	1.3	375	9	AM192111	AM192111	x180B04.x
C 426	20	1.3	375	10	B1299388	B1299388	UI-R-CV2-
C 427	20	1.3	376	10	B1295786	B1295786	UI-R-DK0-
C 428	20	1.3	377	9	BE108471	BE108471	UI-R-CA0-
C 429	20	1.3	378	10	BG668404	BG668404	DRAA0D04
C 430	20	1.3	379	9	AI853989	AI853989	UI-M-BH0-
C 431	20	1.3	379	10	BG629739	BG629739	CC-est1C1
C 432	20	1.3	379	10	BM145111	BM145111	TCAP1D15
C 433	20	1.3	379	10	BE646953	BE646953	UI-M-BH1-
C 434	20	1.3	380	10	BG001344	BG001344	I10-GN028
C 435	20	1.3	381	9	AA899839	AA899839	UI-R-E0-C
C 436	20	1.3	381	9	BE113565	BE113565	UI-R-CA0-
C 437	20	1.3	381	10	BF411276	BF411276	UI-R-B11-
C 438	20	1.3	383	9	AA926019	AA926019	UI-R-Al-e
C 439	20	1.3	383	9	AA961140	AA961140	op14g05.s
C 440	20	1.3	383	10	BG373741	BG373741	UI-R-CV1-
C 441	20	1.3	383	10	BG518547	BG518547	602578283
C 442	20	1.3	384	10	B1305791	B1305791	NL_1.123
C 443	20	1.3	384	10	B1306741	B1306741	NL_5_P06
C 444	20	1.3	386	9	AI556503	AI556503	UI-R-C2P-
C 445	20	1.3	386	9	AA614066	AA614066	no94b12.s
C 446	20	1.3	386	10	BM419968	BM419968	R021E03.O
C 447	20	1.3	389	9	AI556375	AI556375	UI-R-C2P-
C 448	20	1.3	389	9	AA515453	AA515453	ng66606.s
C 449	20	1.3	389	10	BM37629	BM37629	MEST215-B
C 450	20	1.3	391	9	AI219579	AI219579	q905c12.x
C 451	20	1.3	391	10	AM304294	AM304294	xv58d03.x
C 452	20	1.3	391	10	BM383290	BM383290	UI-R-D50-
C 453	20	1.3	391	10	W52121	W52121	zc90601.s1
C 454	20	1.3	391	10	BE511000	BE511000	946057A12
C 455	20	1.3	392	9	AI112848	AI112848	UI-R-E1-f

C 456	20	1.3	393	10	B1703530	B1703530	f691b11.x
C 457	20	1.3	394	9	AI752895	AI752895	cr02a01.x
C 458	20	1.3	395	9	AI849514	AI849514	UI-M-AH1-
C 459	20	1.3	396	9	AI601923	AI601923	UI-R-AD0-
C 460	20	1.3	396	9	AI955893	AI955893	wt31b04.x
C 461	20	1.3	396	10	B1808372	B1808372	C008D11.O
C 462	20	1.3	397	9	AI044054	AI044054	UI-R-Cl-1
C 463	20	1.3	397	9	AI63117	AI63117	tz43e09.x
C 464	20	1.3	397	9	AA488796	AA488796	UI-M-BH3-
C 465	20	1.3	397	10	B1538612	B1538612	434014.MA
C 466	20	1.3	397	12	A2553606	A2553606	RPCI-23-2
C 467	20	1.3	400	9	AV741927	AV741927	AV741927
C 468	20	1.3	400	10	BG042214	BG042214	su93b08.y
C 469	20	1.3	400	10	BF225081	BF225081	uz42g02.y
C 470	20	1.3	402	9	AM015245	AM015245	UI-H-B10P
C 471	20	1.3	402	10	B1142244	B1142244	SD16292.5
C 472	20	1.3	403	9	AL502654	AL502654	AL502654
C 473	20	1.3	404	9	AA614596	AA614596	np50c01.s
C 474	20	1.3	405	10	BE947234	BE947234	UI-M-BH3-
C 475	20	1.3	406	9	BE034134	BE034134	MG05D07.M
C 476	20	1.3	406	9	AA579238	AA579238	nf36c08.s
C 477	20	1.3	406	10	BF788803	BF788803	602110511
C 478	20	1.3	407	9	AA830974	AA830974	oc62b06.s
C 479	20	1.3	407	9	AI502510	AI502510	UI-R-C2-n
C 480	20	1.3	407	9	AI842642	AI842642	UI-M-AQ1-
C 481	20	1.3	407	9	AA489287	AA489287	UI-M-BH3-
C 482	20	1.3	408	12	AA0473338	AA0473338	CITBI-E1-
C 483	20	1.3	409	9	AU031938	AU031938	AU031938
C 484	20	1.3	409	9	AV737699	AV737699	AV737699
C 485	20	1.3	409	9	AM168383	AM168383	x181f05.x
C 486	20	1.3	409	10	BG304056	BG304056	LI29909.x
C 487	20	1.3	409	10	BG379905	BG379905	UI-R-CS0-
C 488	20	1.3	409	10	BG381160	BG381160	UI-R-CS0-
C 489	20	1.3	409	10	BE991431	BE991431	UI-M-B21-
C 490	20	1.3	410	9	AJ281048	AJ281048	4A3A-ABA-
C 491	20	1.3	410	9	AU082987	AU082987	AU082987
C 492	20	1.3	410	10	BE995402	BE995402	UI-M-CG0P
C 493	20	1.3	411	9	AI848379	AI848379	UI-M-AH1-
C 494	20	1.3	411	9	BE115471	BE115471	UI-R-B51-
C 495	20	1.3	411	10	BE947562	BE947562	UI-M-BH3-
C 496	20	1.3	412	9	AI112029	AI112029	UI-R-Y0-m
C 497	20	1.3	412	9	AI837989	AI837989	UI-M-AP0-
C 498	20	1.3	412	10	BM344034	BM344034	IT45h08.y
C 499	20	1.3	413	10	H06065	H06065	y172f11.s1
C 500	20	1.3	414	9	AI840473	AI840473	UI-M-ANO-
C 501	20	1.3	415	9	AA761915	AA761915	n242g11.s
C 502	20	1.3	415	10	BE955667	BE955667	UI-M-BH4-
C 503	20	1.3	416	9	AI791933	AI791933	n166e07.y
C 504	20	1.3	416	9	AI791934	AI791934	n167e07.y
C 505	20	1.3	417	10	BM419932	BM419932	R020H04.O
C 506	20	1.3	417	10	RO6827	RO6827	y111904.r1
C 507	20	1.3	418	9	AU164471	AU164471	AU164471
C 508	20	1.3	419	9	AV564344	AV564344	AV564344
C 509	20	1.3	419	10	B1303736	B1303736	UI-R-DRO-
C 510	20	1.3	420	9	AI555357	AI555357	UI-R-C2P-
C 511	20	1.3	420	9	BB689501	BB689501	BB689501
C 512	20	1.3	421	9	AI732642	AI732642	n167e07.x
C 513	20	1.3	421	9	AU041697	AU041697	AU041697
C 514	20	1.3	421	10	B1277285	B1277285	UI-R-CY0-
C 515	20	1.3	421	12	AQ026290	AQ026290	CITBI-E1-
C 516	20	1.3	422	9	AM135551	AM135551	UI-H-BH1-
C 517	20	1.3	422	10	B1538626	B1538626	434038.MA
C 518	20	1.3	422	12	AQ023926	AQ023926	HS_2259.B
C 519	20	1.3	423	9	AA859669	AA859669	UI-R-E0-b
C 520	20	1.3	423	9	AA578497	AA578497	n166g07.s
C 521	20	1.3	425	9	AA960495	AA960495	vw63c09.s
C 522	20	1.3	425	9	AM522144	AM522144	UI-R-B00-
C 523	20	1.3	425	9	BE111918	BE111918	UI-R-B51-
C 524	20	1.3	426	9	AI384231	AI384231	fd13c11.x
C 525	20	1.3	426	9	AI835212	AI835212	UI-M-AQ0-
C 526	20	1.3	426	12	A2727816	A2727816	RPCI-23-1
C 527	20	1.3	427	9	AI556750	AI556750	UI-R-C2P-
C 528	20	1.3	427	9	AA449306	AA449306	UI-H-B13-

529	20	1.3	427	9	BE023288	sm80d07.y	602	20	1.3	462	10	C98776	C98776	C98776	Rice
C 530	20	1.3	427	10	BE133963	UI-M-BH3	C 603	20	1.3	462	10	BE513153	BE513153	BE513153	UI-H-BW1
C 531	20	1.3	427	10	BE399275	UI-R-CB1	C 604	20	1.3	463	10	BE097186	BE097186	BE097186	UI-R-B01
C 532	20	1.3	428	9	AA765994	oa26h09.s	C 605	20	1.3	463	10	BM219638	BM219638	CO927H12	-
C 533	20	1.3	428	9	AI845539	UI-M-BG0	C 606	20	1.3	463	10	BM219638	BM219638	CO927H12	-
C 534	20	1.3	428	10	BE402885	MI-P-CP1	C 607	20	1.3	464	10	BE248192	BE248192	60185939	-
C 535	20	1.3	428	10	BE716607	NECST3a15	C 608	20	1.3	464	10	BE373188	BE373188	UI-R-BW0	-
C 536	20	1.3	429	9	AI847461	UI-M-A11	C 609	20	1.3	464	10	BM373188	BM373188	UI-R-CV0	-
C 537	20	1.3	429	9	BE117204	UI-R-BB1	C 610	20	1.3	466	10	AI174900	AI174900	AO174900	AO174900
C 538	20	1.3	429	9	BE118965	UI-R-BB1	C 611	20	1.3	466	10	BE105604	BE105604	UI-R-BX0	-
C 539	20	1.3	429	10	BE409214	UI-R-BB1	C 612	20	1.3	466	10	BE457068	BE457068	UI-R-BX0	-
C 540	20	1.3	430	9	AA768669	oa67g12.s	C 613	20	1.3	468	9	AM063504	AM063504	UI-R-BX0	-
C 541	20	1.3	430	9	AV743024	AV743024	C 614	20	1.3	469	9	AI793733	AI793733	UI-R-BX0	-
C 542	20	1.3	431	9	AM355678	pFLC-PK0	C 615	20	1.3	469	9	BB698471	BB698471	UI-R-BX0	-
C 543	20	1.3	431	9	AM874529	hp95d03.x	C 616	20	1.3	470	9	AI840118	AI840118	UI-M-A10	-
C 544	20	1.3	431	9	BE098148	UI-R-BB1	C 617	20	1.3	472	10	AI771691	AI771691	UI-R-Y0-a	-
C 545	20	1.3	431	10	BE291209	602388655	C 618	20	1.3	472	10	BM276299	BM276299	PEFST0aa7	-
C 546	20	1.3	432	9	AM158274	za39c08.x	C 619	20	1.3	473	9	AM253427	AM253427	UI-R-BU0	-
C 547	20	1.3	433	9	AI575420	UI-R-Y0-v	C 620	20	1.3	474	12	AT004712	AT004712	UI-R-BX0	-
C 548	20	1.3	434	9	AI732641	AI732641	C 621	20	1.3	474	12	AZ061856	AZ061856	RPCI-23	-
C 549	20	1.3	434	9	AM958320	EST370390	C 622	20	1.3	475	9	AI840217	AI840217	UI-M-A10	-
C 550	20	1.3	434	12	B49375	RPCI11-6D18	C 623	20	1.3	475	9	AM048091	AM048091	UI-M-A10	-
C 551	20	1.3	435	9	AI850306	UI-M-BG1	C 624	20	1.3	475	9	AM049075	AM049075	UI-M-BH1	-
C 552	20	1.3	435	9	BE096575	UI-R-BU0	C 625	20	1.3	475	9	AM490753	AM490753	UI-M-BH1	-
C 553	20	1.3	435	10	BE993613	BE993613	C 626	20	1.3	476	10	BI703531	BI703531	UI-M-BH3	-
C 554	20	1.3	436	9	AM433868	UI-R-BU0P	C 627	20	1.3	476	9	AA805508	AA805508	oc12e11.s	-
C 555	20	1.3	436	9	BE102585	BE102585	C 628	20	1.3	476	9	BE098524	BE098524	UI-R-BJ1	-
C 556	20	1.3	436	10	BE1759638	BE1759638	C 629	20	1.3	477	10	BC448028	BC448028	UI-R-BJ1	-
C 557	20	1.3	437	9	AI131092	AI131092	C 630	20	1.3	478	10	BE933387	BE933387	UI-M-BZ1	-
C 558	20	1.3	437	10	BI810524	BI810524	C 631	20	1.3	479	10	BE392451	BE392451	UI-R-CB0	-
C 559	20	1.3	438	9	AI840253	UI-M-AH0	C 632	20	1.3	480	9	AI500176	AI500176	UI-R-CB0	-
C 560	20	1.3	438	9	AA171628	AA171628	C 633	20	1.3	480	12	AQ941790	AQ941790	Sheared	-
C 561	20	1.3	438	9	AV736142	AV736142	C 634	20	1.3	481	10	BI354378	BI354378	GM24141.5	-
C 562	20	1.3	438	10	BE377790	BE377790	C 635	20	1.3	483	9	AI840143	AI840143	UI-M-AH0	-
C 563	20	1.3	438	10	CI8256	CI8256	C 636	20	1.3	483	9	AU094483	AU094483	UI-M-AH0	-
C 564	20	1.3	439	9	AI201374	AI201374	C 637	20	1.3	484	10	BE110465	BE110465	7n37608.x	-
C 565	20	1.3	439	12	AQ406831	AQ406831	C 638	20	1.3	485	10	BE512505	BE512505	UI-R-BW1	-
C 566	20	1.3	440	10	BE113876	BE113876	C 639	20	1.3	486	12	AZ616655	AZ616655	1M0446N04	-
C 567	20	1.3	440	10	BI302654	BI302654	C 640	20	1.3	486	12	A2852467	A2852467	2M0155003	-
C 568	20	1.3	442	9	AI167706	AI167706	C 641	20	1.3	487	9	AI839728	AI839728	UI-M-AH0	-
C 569	20	1.3	442	10	BE242565	BE242565	C 642	20	1.3	487	9	AI958638	AI958638	fd12c01.x	-
C 570	20	1.3	442	10	BE714603	BE714603	C 643	20	1.3	487	10	BI998324	BI998324	1031056CQ	-
C 571	20	1.3	443	9	AI1713065	AI1713065	C 644	20	1.3	488	9	AI521472	AI521472	th61d12.x	-
C 572	20	1.3	443	9	AM297806	AM297806	C 645	20	1.3	488	10	BE388529	BE388529	UI-R-CB1	-
C 573	20	1.3	443	9	AM613525	AM613525	C 646	20	1.3	490	10	BI316068	BI316068	saF59d12	-
C 574	20	1.3	445	9	AI182617	AI182617	C 647	20	1.3	490	10	BM378910	BM378910	MST571-C	-
C 575	20	1.3	445	9	AV713159	AV713159	C 648	20	1.3	492	12	AQ080162	AQ080162	CIT-HSP-2	-
C 576	20	1.3	446	9	AL640817	AL640817	C 649	20	1.3	493	9	AM740688	AM740688	uf40h04.y	-
C 577	20	1.3	446	9	AM487835	AM487835	C 650	20	1.3	494	10	BI297999	BI297999	UI-R-CV2	-
C 578	20	1.3	447	10	BE964136	BE964136	C 651	20	1.3	495	9	AI602504	AI602504	UI-R-AG0	-
C 579	20	1.3	447	9	AI045361	AI045361	C 652	20	1.3	496	10	BE542674	BE542674	UI-R-C3-S	-
C 580	20	1.3	447	9	AI172588	AI172588	C 653	20	1.3	498	9	AI146046	AI146046	UI-R-BT0	-
C 581	20	1.3	449	9	AI835687	AI835687	C 654	20	1.3	499	10	BE712771	BE712771	MT-P-H3-a	-
C 582	20	1.3	449	9	AI0184281	AI0184281	C 655	20	1.3	500	9	AM532067	AM532067	UI-R-BU0	-
C 583	20	1.3	451	9	AI060124	AI060124	C 656	20	1.3	500	10	BE416926	BE416926	UI-R-CN0	-
C 584	20	1.3	451	9	AI753272	AI753272	C 657	20	1.3	501	9	AI836244	AI836244	UI-M-AH0	-
C 585	20	1.3	451	10	BM080330	BM080330	C 658	20	1.3	503	9	BE275086	BE275086	AV565629	-
C 586	20	1.3	451	10	R42740	R42740	C 659	20	1.3	504	9	AL372177	AL372177	CA_EB002	-
C 587	20	1.3	452	9	AA807990	AA807990	C 660	20	1.3	504	10	BE842411	BE842411	MEST30-El0	-
C 588	20	1.3	452	9	AI059226	AI059226	C 661	20	1.3	504	10	BE507554	BE507554	UI-H-BW1	-
C 589	20	1.3	453	9	AV741708	AV741708	C 662	20	1.3	505	9	AA073291	AA073291	mm95c03.r	-
C 590	20	1.3	454	9	AI687558	tp97a10.x	C 663	20	1.3	505	9	AI521184	AI521184	tl58e04.x	-
C 591	20	1.3	454	9	AV540583	AV540583	C 664	20	1.3	506	9	AV737170	AV737170	AV737170	-
C 592	20	1.3	456	9	AI1444896	AI1444896	C 665	20	1.3	506	10	BE392714	BE392714	UI-R-CA0	-
C 593	20	1.3	456	9	AM298816	AM298816	C 666	20	1.3	507	9	AI501434	AI501434	UI-R-C2P	-
C 594	20	1.3	458	9	AI299084	AI299084	C 667	20	1.3	507	9	AM329803	AM329803	N201075e	-
C 595	20	1.3	458	10	BM223942	BM223942	C 668	20	1.3	508	9	AI716615	AI716615	UI-R-Y0-a	-
C 596	20	1.3	458	10	BM313642	BM313642	C 669	20	1.3	508	10	BI280347	BI280347	UI-R-DE0	-
C 597	20	1.3	459	9	AM134739	AM134739	C 670	20	1.3	508	10	BM387746	BM387746	UI-R-CN1	-
C 598	20	1.3	459	10	BM276216	BM276216	C 671	20	1.3	509	9	AI395770	AI395770	MM005120	-
C 599	20	1.3	462	9	AI502077	AI502077	C 672	20	1.3	509	9	AL370508	AL370508	MCA38C04	-
C 600	20	1.3	462	10	BM381987	BM381987	C 673	20	1.3	509	10	BI281413	BI281413	UI-R-DH0	-
C 601	20	1.3	462	10	C22351	C22351	C 674	20	1.3	509	10	BI801132	BI801132	H040C09	-

C 675	20	1.3	510	9	AM138907	UI-H-B11-	748	20	1.3	554	9	AF201273	AF201273
C 676	20	1.3	510	10	BE949082	UI-M-BH3-	C 749	20	1.3	554	9	AM290990	AM290990
C 677	20	1.3	512	9	AI138056	UI-R-C2P-	C 750	20	1.3	554	10	BE678064	BE678064
C 678	20	1.3	512	9	AM534385	UI-R-C4-a	C 751	20	1.3	555	9	BE097003	BE097003
C 679	20	1.3	512	9	AM539506	C0120F09-	C 752	20	1.3	557	10	BM154097	BM154097
C 680	20	1.3	512	10	BG381422	UI-R-CT0-	C 753	20	1.3	557	9	AA875069	AA875069
C 681	20	1.3	513	9	AI836243	UI-M-AP0-	C 754	20	1.3	559	10	BG063777	BG063777
C 682	20	1.3	513	9	AI970339	UI-R-C240	C 755	20	1.3	559	10	B1275694	B1275694
C 683	20	1.3	513	10	BG380507	UI-R-CT0-	C 756	20	1.3	560	10	B1310016	B1310016
C 684	20	1.3	515	9	AM296931	UI-H-BM0-	C 757	20	1.3	560	10	B1886843	B1886843
C 685	20	1.3	515	10	BE950659	UI-M-CE0-	C 758	20	1.3	561	10	BM336250	BM336250
C 686	20	1.3	516	9	AJ280858	AA3A-AAI-	C 759	20	1.3	562	9	AM137497	AM137497
C 687	20	1.3	516	9	AM060450	UI-M-BH1-	C 760	20	1.3	565	10	BM274784	BM274784
C 688	20	1.3	516	10	BI067754	pglin.-PK0	C 761	20	1.3	565	10	BF507500	BF507500
C 689	20	1.3	518	10	BE466475	UI-M-CP0	C 762	20	1.3	566	9	AV533051	AV533051
C 690	20	1.3	520	9	AA761138	AA761138	C 763	20	1.3	566	10	BF171742	BF171742
C 691	20	1.3	520	9	AM270042	xv57f08.x	C 764	20	1.3	569	10	BE548173	BE548173
C 692	20	1.3	521	9	AM281944	fj59e04.x	C 765	20	1.3	569	12	A2967841	A2967841
C 693	20	1.3	521	10	BF403788	UI-R-CA1-	C 766	20	1.3	570	10	BM274795	BM274795
C 694	20	1.3	522	9	AI714505	UI-R-CA1-	C 767	20	1.3	572	12	A2659203	A2659203
C 695	20	1.3	522	10	BI161850	ME9170-C	C 768	20	1.3	574	10	B1721485	B1721485
C 696	20	1.3	522	10	BF196350	7n68g09.x	C 769	20	1.3	575	9	AV715655	AV715655
C 697	20	1.3	523	9	AI172538	UI-R-C2P-	C 770	20	1.3	575	9	AA453166	AA453166
C 698	20	1.3	523	10	BG381203	UI-R-CT0-	C 771	20	1.3	576	9	AI714497	AI714497
C 699	20	1.3	524	9	AI713300	UI-R-ABI-	C 772	20	1.3	577	12	AQ410094	AQ410094
C 700	20	1.3	524	9	AV736529	AV736529	C 773	20	1.3	578	12	A2458849	A2458849
C 701	20	1.3	524	9	AM449230	UI-H-B13-	C 774	20	1.3	579	9	AI714504	AI714504
C 702	20	1.3	524	10	BM388489	UI-R-DE0-	C 775	20	1.3	581	10	BG487376	BG487376
C 703	20	1.3	526	10	BI227170	UI-R-CY0-	C 776	20	1.3	582	10	BG835077	BG835077
C 704	20	1.3	526	10	BF400655	UI-R-CA0-	C 777	20	1.3	585	9	BE111753	BE111753
C 705	20	1.3	527	10	BG381438	UI-R-CT0-	C 778	20	1.3	585	10	BI404546	BI404546
C 706	20	1.3	527	10	BI067769	pglin.-PK0	C 779	20	1.3	585	10	BF408909	BF408909
C 707	20	1.3	527	10	BM332534	ME9170-C	C 780	20	1.3	587	9	AM706930	AM706930
C 708	20	1.3	527	10	BF012490	ux57f05.Y	C 781	20	1.3	587	9	AM970928	AM970928
C 709	20	1.3	528	12	AQ453166	HS.5194_A	C 782	20	1.3	588	9	AV739462	AV739462
C 710	20	1.3	529	9	AA939218	CP74M09.X	C 783	20	1.3	589	9	AA579787	AA579787
C 711	20	1.3	529	9	AI057550	OY30M04.X	C 784	20	1.3	589	9	AI714516	AI714516
C 712	20	1.3	529	9	AM049672	UI-M-BH1-	C 785	20	1.3	590	10	AI013558	AI013558
C 713	20	1.3	529	9	AM119986	707006F09	C 786	20	1.3	590	10	BG807236	BG807236
C 714	20	1.3	529	10	BM379950	ME9170-C	C 787	20	1.3	591	10	BI428793	BI428793
C 715	20	1.3	530	9	AI826069	wk28B10.X	C 788	20	1.3	591	10	BE511297	BE511297
C 716	20	1.3	530	9	AM444929	UI-H-B13-	C 789	20	1.3	594	10	BI805257	BI805257
C 717	20	1.3	530	10	BF404279	UI-R-CA1-	C 790	20	1.3	598	10	BM339041	BM339041
C 718	20	1.3	530	12	BH061449	RPCT-24-3	C 791	20	1.3	599	10	BM347917	BM347917
C 719	20	1.3	531	9	AU108602	AU108602	C 792	20	1.3	599	10	BF420639	BF420639
C 720	20	1.3	531	9	AI112963	UI-R-C0-9	C 793	20	1.3	600	10	BG802901	BG802901
C 721	20	1.3	533	9	AV754613	AV754613	C 794	20	1.3	600	10	BG835078	BG835078
C 722	20	1.3	533	10	BE440045	HTM1-816F	C 795	20	1.3	600	10	BM382367	BM382367
C 723	20	1.3	534	9	AI834736	606072R04	C 796	20	1.3	604	9	AI982412	AI982412
C 724	20	1.3	534	9	BE946259	UI-M-BH3-	C 797	20	1.3	604	10	BE509219	BE509219
C 725	20	1.3	534	10	BE669073	DRN05C05	C 798	20	1.3	606	9	BE098809	BE098809
C 726	20	1.3	534	12	AQ917303	T233284B	C 799	20	1.3	606	9	AL585481	AL585481
C 727	20	1.3	535	10	BE651817	UI-M-AP0-	C 800	20	1.3	607	9	AM205382	AM205382
C 728	20	1.3	536	9	AM135047	UI-H-B11-	C 801	20	1.3	610	9	BF398606	BF398606
C 729	20	1.3	540	10	BI344041	372391.MA	C 802	20	1.3	610	10	BE681924	BE681924
C 730	20	1.3	541	9	AI841813	UI-M-AN0-	C 803	20	1.3	612	10	BM276497	BM276497
C 731	20	1.3	541	10	BE669073	DRN05C05	C 804	20	1.3	612	12	AQ972343	AQ972343
C 732	20	1.3	542	9	AI575029	UI-R-G0-U	C 805	20	1.3	613	9	AI714498	AI714498
C 733	20	1.3	542	10	BG379482	UI-R-CS0-	C 806	20	1.3	613	10	BE087268	BE087268
C 734	20	1.3	544	9	AM296798	UI-H-BM0-	C 807	20	1.3	615	9	AL500113	AL500113
C 735	20	1.3	544	10	BM384134	UI-R-D20-	C 808	20	1.3	615	12	AG053783	AG053783
C 736	20	1.3	546	10	BE673578	DRN03G03	C 809	20	1.3	618	10	BG379351	BG379351
C 737	20	1.3	547	9	AM251158	UI-R-BJ0-	C 810	20	1.3	619	9	AM091504	AM091504
C 738	20	1.3	548	9	AL584003	AL584003	C 811	20	1.3	626	10	BM386877	BM386877
C 739	20	1.3	549	9	AV543498	AV543498	C 812	20	1.3	632	9	AI714874	AI714874
C 740	20	1.3	549	9	AM531891	UI-R-BSD-	C 813	20	1.3	636	9	BE108211	BE108211
C 741	20	1.3	549	10	BM418528	R001A02.O	C 814	20	1.3	636	10	BE110554	BE110554
C 742	20	1.3	550	10	BF393807	UI-R-CA0-	C 815	20	1.3	645	9	BE110554	BE110554
C 743	20	1.3	551	9	AA859508	UI-R-E0-b	C 816	20	1.3	646	10	BE1562546	BE1562546
C 744	20	1.3	551	9	AA201992	ID05062.5	C 817	20	1.3	647	10	BE680287	BE680287
C 745	20	1.3	552	9	AM154085	f122c07.x	C 818	20	1.3	648	9	AV724415	AV724415
C 746	20	1.3	553	10	BM274623	PIEST0a44	C 819	20	1.3	650	10	C97989	C97989
C 747	20	1.3	554	9	AF201193	AF201193	C 820	20	1.3				

821	20	1.3	651	10	BE312846	601146107	894	20	1.3	839	10	BF540920
822	20	1.3	652	10	BI390392	BI390392	895	20	1.3	840	12	A0751817
823	20	1.3	653	10	BI312259	EST531400	896	20	1.3	841	11	AF078845
824	20	1.3	654	10	BF297032	046PXC05	897	20	1.3	845	10	BG387410
825	20	1.3	655	9	AL642590	AL642590	898	20	1.3	845	10	BE899369
826	20	1.3	656	9	AL036786	AL036786	899	20	1.3	847	10	BI756798
827	20	1.3	657	10	W48389	W48389	900	20	1.3	851	10	BF267529
828	20	1.3	657	12	BH064002	BH064002	901	20	1.3	851	12	BH132843
829	20	1.3	661	10	BM268717	BM268717	902	20	1.3	857	12	CNS01H7X
830	20	1.3	663	9	BB641337	BB641337	903	20	1.3	854	10	BG110499
831	20	1.3	663	10	BF404565	BF404565	904	20	1.3	875	10	BG367233
832	20	1.3	664	9	AA753540	AA753540	905	20	1.3	876	10	BF343974
833	20	1.3	664	12	BH454698	BH454698	906	20	1.3	892	12	AC080722
834	20	1.3	665	10	BF240191	BF240191	907	20	1.3	894	10	BF671660
835	20	1.3	667	10	BM384384	BM384384	908	20	1.3	917	10	BF179155
836	20	1.3	671	10	BE949481	BE949481	909	20	1.3	924	10	BF123125
837	20	1.3	673	10	BB221493	BB221493	910	20	1.3	939	11	BC014503
838	20	1.3	673	10	BG840789	BG840789	911	20	1.3	939	12	A0896169
839	20	1.3	681	10	BF724733	BF724733	912	20	1.3	984	12	CNS037OR
840	20	1.3	682	12	BH003388	BH003388	913	20	1.3	989	9	AI964330
841	20	1.3	684	12	BH002809	BH002809	914	20	1.3	1034	10	BG622094
842	20	1.3	687	10	AZ339135	AZ339135	915	20	1.3	1111	11	BC015102
843	20	1.3	699	9	BG352215	BG352215	916	20	1.3	1121	10	BI834141
844	20	1.3	699	9	AU161189	AU161189	917	20	1.3	1314	10	BM476808
845	20	1.3	699	9	BF165989	BF165989	918	20	1.3	1436	10	BF797425
846	20	1.3	701	12	AZ647643	AZ647643	919	20	1.3	1520	11	AK004673
847	20	1.3	703	10	BE966414	BE966414	920	20	1.3	1721	11	BC020285
848	20	1.3	704	10	BE612997	BE612997	921	20	1.3	1733	11	BC012859
849	20	1.3	705	12	BH553303	BH553303	922	20	1.3	2236	11	BC009185
850	20	1.3	709	9	BB309095	BB309095	923	20	1.3	23	2	HSM002743
851	20	1.3	711	10	C97976	C97976	924	20	1.2	23	2	HSM002743
852	20	1.3	714	9	AL048406	AL048406	925	19	1.2	23	2	HSM003068
853	20	1.3	717	12	A0888285	A0888285	926	19	1.2	23	2	HSM003085
854	20	1.3	718	9	AV713508	AV713508	927	19	1.2	23	2	HSM003164
855	20	1.3	722	9	BE035676	BE035676	928	19	1.2	30	2	HSM001042
856	20	1.3	722	10	BI279113	BI279113	929	19	1.2	32	2	HSM001595
857	20	1.3	723	10	BI838318	BI838318	930	19	1.2	32	2	HSM001674
858	20	1.3	724	10	BF624715	BF624715	931	19	1.2	33	2	HSM003086
859	20	1.3	725	10	BE279434	BE279434	932	19	1.2	33	2	HSM003165
860	20	1.3	728	10	BG273735	BG273735	933	19	1.2	33	9	AL048719
861	20	1.3	729	10	BG088014	BG088014	934	19	1.2	33	9	AL048733
862	20	1.3	729	10	BG344353	BG344353	935	19	1.2	34	2	HSM003069
863	20	1.3	731	12	BH502887	BH502887	936	19	1.2	49	10	BG361220
864	20	1.3	732	10	BG611288	BG611288	937	19	1.2	51	9	AA589046
865	20	1.3	734	9	AL048393	AL048393	938	19	1.2	53	2	HSM002757
866	20	1.3	736	10	BI294856	BI294856	939	19	1.2	53	10	BM142044
867	20	1.3	738	10	BM389512	BM389512	940	19	1.2	56	10	BI502320
868	20	1.3	738	10	BE881237	BE881237	941	19	1.2	59	10	BI742219
869	20	1.3	738	12	AQ113924	AQ113924	942	19	1.2	63	10	BG361727
870	20	1.3	747	9	AM963627	AM963627	943	19	1.2	63	10	BG962258
871	20	1.3	757	10	BI860173	BI860173	944	19	1.2	65	2	HSM001786
872	20	1.3	758	12	BH248122	BH248122	945	19	1.2	66	9	AI348706
873	20	1.3	762	9	AL535104	AL535104	946	19	1.2	66	9	AA572834
874	20	1.3	763	10	BI254402	BI254402	947	19	1.2	66	10	BG361618
875	20	1.3	766	10	BE966692	BE966692	948	19	1.2	66	12	CNS02DN9
876	20	1.3	770	9	AL584124	AL584124	949	19	1.2	67	9	AI271601
877	20	1.3	777	10	BI078603	BI078603	950	19	1.2	68	10	BG113507
878	20	1.3	778	12	AG095725	AG095725	951	19	1.2	70	9	AI309514
879	20	1.3	779	10	BE568925	BE568925	952	19	1.2	71	10	BM026556
880	20	1.3	787	9	AA732467	AA732467	953	19	1.2	74	9	AI683253
881	20	1.3	787	10	BG674589	BG674589	954	19	1.2	76	9	AI309824
882	20	1.3	789	12	AZ981604	AZ981604	955	19	1.2	77	9	AA980014
883	20	1.3	790	12	BH567699	BH567699	956	19	1.2	80	9	AA009051
884	20	1.3	790	12	BH455256	BH455256	957	19	1.2	81	10	N82562
885	20	1.3	795	12	BH455256	BH455256	958	19	1.2	81	10	T25577
886	20	1.3	799	10	BG740298	BG740298	959	19	1.2	83	10	BG361896
887	20	1.3	799	12	AQ751130	AQ751130	960	19	1.2	83	10	BG362080
888	20	1.3	805	9	AU117404	AU117404	961	19	1.2	85	10	BG792352
889	20	1.3	809	10	BI651159	BI651159	962	19	1.2	85	10	BG817145
890	20	1.3	810	10	BG442801	BG442801	963	19	1.2	86	10	BM026438
891	20	1.3	827	10	BF984513	BF984513	964	19	1.2	88	9	AA582651
892	20	1.3	838	10	BF264841	BF264841	965	19	1.2	88	10	BE662018
893	20	1.3	838	12	AZ201255	AZ201255	966	19	1.2	89	9	AI223740
894	20	1.3	839	10	BF540920	BF540920	894	20	1.3	839	10	BF540920
895	20	1.3	840	12	A0751817	A0751817	895	20	1.3	840	12	A0751817
896	20	1.3	841	11	AF078845	AF078845	896	20	1.3	841	11	AF078845
897	20	1.3	845	10	BG387410	BG387410	897	20	1.3	845	10	BG387410
898	20	1.3	845	10	BE899369	BE899369	898	20	1.3	845	10	BE899369
899	20	1.3	847	10	BI756798	BI756798	899	20	1.3	847	10	BI756798
900	20	1.3	851	10	BF267529	BF267529	900	20	1.3	851	10	BF267529
901	20	1.3	851	12	BH132843	BH132843	901	20	1.3	851	12	BH132843
902	20	1.3	857	12	CNS01H7X	CNS01H7X	902	20	1.3	857	12	CNS01H7X
903	20	1.3	854	10	BG110499	BG110499	903	20	1.3	854	10	BG110499
904	20	1.3	875	10	BG367233	BG367233	904	20	1.3	875	10	BG367233
905	20	1.3	876	10	BF343974	BF343974	905	20	1.3	876	10	BF343974
906	20	1.3	892	12	AC080722	AC080722	906	20	1.3	892	12	AC080722
907	20	1.3	894	10	BF671660	BF671660	907	20	1.3	894	10	BF671660
908	20	1.3	917	10	BF179155	BF179155	908	20	1.3	917	10	BF179155
909	20	1.3	924	10	BF123125	BF123125	909	20	1.3	924	10	BF123125
910	20	1.3	939	11	BC014503	BC014503	910	20	1.3	939	11	BC014503
911	20	1.3	939	12	A0896169	A0896169	911	20	1.3	939	12	A0896169
912	20	1.3	984	12	CNS037OR	CNS037OR	912	20	1.3	984	12	CNS037OR
913	20	1.3	989	9	AI964330	AI964330	913	20	1.3	989	9	AI964330
914	20	1.3	1034	10	BG622094	BG622094	914	20	1.3	1034	10	BG622094
915	20	1.3	1111	11	BC015102	BC015102	915	20	1.3	1111	11	BC015102
916	20	1.3	1121	10	BI834141	BI834141	916	20	1.3	1121	10	BI834141
917	20	1.3	1314	10	BM476808	BM476808	917	20	1.3	1314	10	BM476808
918	20	1.3	1436	10	BF797425	BF797425	918	20	1.3	1436	10	BF797425
919	20	1.3	1520	11	AK004673	AK004673	919	20	1.3	1520	11	AK004673
920	20	1.3	1721	11	BC020285	BC020285	920	20	1.3	1721	11	BC020285
921	20	1.3	1733	11	BC012859	BC012859	921	20	1.3	1733	11	BC012859
922	20	1.3	2236	11	BC009185	BC009185	922	20	1.3	2236	11	BC009185
923	20	1.3	23	2	HSM002743	HSM002743	923	20	1.3	23	2	HSM002743
924	20	1.2	23	2	HSM003068	HSM003068	924	20	1.2	23	2	HSM003068
925	19	1.2	23	2	HSM003085	HSM003085	925	19	1.2	23	2	HSM003085
926												

```

967 19 1.2 89 12 CNS04J97 AL293236 Tetraodon
968 19 1.2 94 9 AJ318247 AJ318247
969 19 1.2 99 2 HSM001797 AJ103746 Homo sapi
970 19 1.2 100 9 AI310259 AI310259
971 19 1.2 100 9 AU038797 AU038797
972 19 1.2 101 9 AI252895 AI252895
973 19 1.2 103 9 AI431004 AI431004
974 19 1.2 104 9 AI250713 AI250713
975 19 1.2 104 9 AJ318296 AJ318296
976 19 1.2 104 9 AW336521 AW336521
977 19 1.2 105 9 BI470550 BI470550
978 19 1.2 105 9 AW415490 AW415490
979 19 1.2 105 9 BG360183 BG360183
980 19 1.2 106 9 AW262956 AW262956
981 19 1.2 106 9 BG153744 BG153744
982 19 1.2 107 9 BG360957 BG360957
983 19 1.2 111 9 BM161838 BM161838
984 19 1.2 112 9 BF883795 BF883795
985 19 1.2 112 9 A2227488 A2227488
986 19 1.2 114 9 BG151216 BG151216
987 19 1.2 115 9 BF415632 BF415632
988 19 1.2 117 9 BE043260 BE043260
989 19 1.2 119 9 AW484731 AW484731
990 19 1.2 119 9 BE987817 BE987817
991 19 1.2 121 9 BG362335 BG362335
992 19 1.2 122 9 BI280842 BI280842
993 19 1.2 123 9 BI399017 BI399017
994 19 1.2 124 9 AW091178 AW091178
995 19 1.2 124 9 BG913282 BG913282
996 19 1.2 124 9 BE757520 BE757520
997 19 1.2 125 9 AA769981 AA769981
998 19 1.2 125 9 AU165528 AU165528
999 19 1.2 128 9 AI763694 AI763694
1000 19 1.2 128 9 AL506504 AL506504

```

ALIGNMENTS

```

RESULT 1
LOCUS AI542842 448 bp mRNA linear EST 23-APR-2001
DEFINITION SD09326.5prime SD Drosophila melanogaster Schneider L2 cell culture
port2 Drosophila melanogaster cDNA clone SD09326 5 similar to
CG11822: FBan0011822 'ion channel' located on: 2L 21C5-21C5;:
04/13/2001, mRNA sequence.
ACCESSION AI542842 GI:13771762
VERSION 1
KEYWORDS EST.
SOURCE fruit fly.
ORGANISM Drosophila melanogaster
Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
Pterygota; Neoptera; Endopterygota; Diptera; Brachycera;
Muscomorpha; Ephydroidea; Drosophilidae; Drosophila.
REFERENCE 1 (bases 1 to 448)
AUTHORS Harvey,D., Brokstein,P., Hong,L., Evans-Holm,M., Su,C., Tsang,G.,
Lewis,S. and Rubin,G.M.
TITLE BDGP/HMI Drosophila EST Project
JOURNAL Unpublished (2001)
COMMENT On Mar 19, 1999 this sequence version replaced gi:4460215.
Other_ESTs: SD09326.3prime
Contact: Stapleton, M.
BDGP Lawrence Berkeley National Lab
One Cyclotron Rd Berkeley, CA 94720, USA
Fax: 510 486 6798
Email: http://www.fruitfly.org/EST, est@fruitfly.berkeley.edu
hit genomic AB003589: arm:2L [301580,604052]
estimated-cyto:21B7-21C7: 04/13/2001
Plate: SD.93 row: C column: 2
High quality sequence stop: 366
POLYA-No.
Location/Qualifiers

```

FEATURES

```

source 1. 448
/organism="Drosophila melanogaster"
/db_xref="taxon:7227"
/clone="SD09326"
/clone_lib="SD Drosophila melanogaster Schneider L2 cell
culture port2"
/lab_host="DH5-alpha"
/note="Vector: port2; Site_1: EcoRI, site_2: XhoI; Sized
fractionated cDNAs were directly ligated into port2.
Plasmid cDNA library."
BASE COUNT 115 a 140 c 116 g 77 t
ORIGIN

```

```

Query Match 18.7%; Score 288; DB 9; Length 448;
Best local Similarity 99.3%; Pred. No. 1.9e-73;
Matches 438; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

```

```

QY 14 gtacatcgaacaaagcgcgctgaacaatgaagacgactcccaagataaagccag 73
DB 8 gTACATCCGAACAAAGCGCGCTGAACATGACGACGACTCCCAAGATAAAGGCACAG 67
QY 74 ttccggtcctgagtcgcaactactctgcaaatgtatgggagatgtcttatgggc 133
DB 68 TTTCCGGTCTCGACTGCGCACTACTCTGCAAAATGCTTAATGGGAGCTTTATGGGCG 127
QY 134 tgacttcgtgcaagcgccgcaactgcaacgcaagcgaagcgaatgtcaagccg 193
DB 128 TGACTTCGTGCAAGCGCGCACTGCAACGCGGACCCCAAGAGCAATGTCAAGGGCG 187
QY 194 tggatcgctcgaagcgccgctgttcaagaaatgaagcagcgatgtgcagcggttcc 253
DB 188 TGATCGCTCGCAAGCGCGCGCTGTTCAGCAACATGACAGCGATGTGACGCGGTTCG 247
QY 254 aaggaaccccccaagcagcgttcctggaatggtgttcactaagaagcatcgaagat 313
DB 248 AAGGAACCCCAAGCAAGCTGCTCCGTGGAATGTTGTCTCACTATGACATGACATGACAGAT 307
QY 314 tgaacgcaagctgaccaccacactgctgctgatactccgatgagagacgaagcgcg 373
DB 308 TGAACGCAAGCTGACCACCACTGCTGCTGATTCGATGAGAGAGAGAGACGGCG 367
QY 374 tgtgcaacgctcaaatatgacaacatcaacgacatcaattgaagtcagcgaggtc 433
DB 368 TGTGCAACCGCTCACTATGACATGACATGACATGACATGACATGACATGACATGACATG 427
QY 434 ggaaccccccaatcacactc 454
DB 428 GGACCCCGCAATCACACTCT 448

```

```

RESULT 2
LOCUS CNS013TY/C 890 bp DNA linear GSS 28-JUL-1999
DEFINITION Drosophila melanogaster genome survey sequence Sp6 end of BAC
BACN10F18 of DrosBAC library from Drosophila melanogaster (fruit
fly), genomic survey sequence.
ACCESSION AL103312 GI:5614923
VERSION 1
KEYWORDS GSS.
SOURCE fruit fly.
ORGANISM Drosophila melanogaster
Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
Pterygota; Neoptera; Endopterygota; Diptera; Brachycera;
Muscomorpha; Ephydroidea; Drosophilidae; Drosophila.
REFERENCE 1 (bases 1 to 890)
AUTHORS Genoscope.
TITLE Direct Submissiion
JOURNAL Submitted (23-JUL-1999) Genoscope - Centre National de Sequencage :
BP 191 91006 EVRY cedex - FRANCE (E-mail : sequef@genoscope.cns.fr)
- Web : www.genoscope.cns.fr)
COMMENT determination of this BAC-end sequence was carried out as part of a
collaboration with the European Drosophila Genome Project (BDGP) -

```

http://www.edgp.ebi.ac.uk . This Drosophila melanogaster BAC library (Dros BAC) was made by Alain Billaud at CEPH (Centre d'Etude du Polymorphisme Humain) with funding provided by a MRC project grant. The DNA was prepared from embryos by Alain Bucheton and Genevieve Payan. It has been constructed in the vector pBelobAC11.

FEATURES
source
1. 890
/organism="Drosophila melanogaster"
/plasmid="pBelobAC11"
/db_xref="taxon:7227"
/clone_lib="DrosBAC"
/clone="BACN10F18"
/note="end : SP6"

BASE COUNT
ORIGIN
205 a 193 c 240 g 198 t 54 others

Query Match
Best Local Similarity 100.0%; Score 162; DB 12; Length 890;
Matches 162; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 568 ctcaatgagcccaagcaagcagactgcaagtgaagatcgctcctgggacctgaag 627
DB 310 CTCACTGGCCCGCCACGACAGCAGCTGCAAGTTGAAGATCGGCTCGGCGCTGAAG 251
QY 628 gtgctctgcgcgagacgagcagcgagagagatccttcgacacgagacctgctt 687
DB 250 GTCGCTCTGCGGAGAACGGACGCGAGAGAGAGTCCCTTGACCAAGAGACTGTT 191
QY 688 cagtcaccgagtggaatcgtagctcgagaccacttt 729
DB 190 CAGTCACCGGAGTGGAATCGTGTGACTCGGAGCCCACTTT 149

RESULT 3
AM944588/c
LOCUS
DEFINITION
361 bp mRNA linear EST 05-JUL-2001
SD09326.3prime SD Drosophila melanogaster Schneider L2 cell culture
POT2 Drosophila melanogaster cDNA clone SD09326.3 similar to
AC004573: Drosophila melanogaster, chromosome 2L, region 21C5-21D1,
P1 clone DS07610, complete sequence, mRNA sequence.

ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
AM944588.1 GI:8122336
EST

fruit fly.
Drosophila melanogaster
Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
Pterygota; Neoptera; Endopterygota; Diptera; Brachycera;
Muscomorpha; Ephydroidea; Drosophilidae; Drosophila.
1 (bases 1 to 361)

REFERENCE
AUTHORS
TITLE
JOURNAL
COMMENT
Harvey, D., Brokstein, P., Hong, L., Evans-Holm, M., Su, C., Tsang, G.,
Lewis, S. and Rubin, G.M.
BDGP/HMI Drosophila EST Project
Unpublished (2001)
Other ESTs: SD09326.5prime
Contact: Stapleton, M.

BDGP
Lawrence Berkeley National Lab
One Cyclotron Rd, Berkeley, CA 94720, USA
Fax: 510 486 6798

Email: http://www.fruitfly.org/EST, est@fruitfly.berkeley.edu
Based upon the presence of a XhoI site followed by a run of 14 or
more T residues at the beginning of the sequence, this clone was
polydenylated. The resulting poly-T sequence has been removed. hit
genomic AE003589: Drosophila melanogaster genomic scaffold
14200013386046 section 15 of 16, complete sequence.: 03/19/2001
Plate: SD.93 row: C column: 2
High quality sequence stop: 281.

FEATURES
source
1. 361
/organism="Drosophila melanogaster"
/db_xref="taxon:7227"

BASE COUNT
ORIGIN
80 a 98 c 102 g 80 t 1 others

Query Match
Best Local Similarity 100.0%; Score 57; DB 9; Length 361;
Matches 57; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1413 tgaactagtttgattcgattcattcatgtatttaattgtgtgagactataatt 1469
DB 99 TGAAC TAGTTTGCATTTGCATTTGCATTTGAATGTTGTCGCAACTATTAATT 43

RESULT 4
BI285394/c
LOCUS
DEFINITION
303 bp mRNA linear EST 19-JUL-2001
UI-R-DB0-byz-e-12-0-UI.s1 UI-R-DB0 Rattus norvegicus cDNA clone
UI-R-DB0-byz-e-12-0-UI 3', mRNA sequence.
BI285394
ACCESSION
VERSION
KEYWORDS
SOURCE
EST.
Norway rat.
Rattus norvegicus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
Rattus.

REFERENCE
AUTHORS
TITLE
JOURNAL
MEDLINE
COMMENT
1 (bases 1 to 303)
Bonaldo, M.F., Lennon, G. and Soares, M.B.
Normalization and subtraction: two approaches to facilitate gene
discovery
Genome Res. 6 (9), 791-806 (1996)
97044477
Contact: Soares, MB
Program for Rat Gene Discovery and Mapping
University of Iowa
451 Eckstein Medical Research Building Iowa City, IA 52242, USA
Tel: 319 335 8250
Fax: 319 335 9565

Email: msoares@blue.weeg.uiowa.edu
The sequence contained an oligo-dT track that was present in the
oligonucleotide that was used to prime the synthesis of first
strand cDNA and therefore this may represent a bonafide poly A
tail. The sequence tag present in the cDNA between the NotI site
and the oligo-dT track served to verify it as a clone from the
non-normalized bladder library cDNA Library Preparation: M.B.
Soares Lab Clone distribution: clones will be available through
Research Genetics (www.resgen.com)
Seq primer: M13 forward
POLYA-yes.

FEATURES
source
1. 303
Location/Qualifiers

/organism="Rattus norvegicus"
/strain="Sprague-Dawley"
/db_xref="taxon:10116"
/clone="UI-R-DB0-byz-e-12-0-UI"
/clone_lib="UI-R-DB0"
/dev_stage="ADULT"
/lab_host="DH10B (Life Technologies)"
/note="Vector: pT73D-Pac (Pharmacia) with a modified
polylinker. Site 1: Not I; Site 2: Eco RI; The UI-R-DB0
library is a non-normalized library constructed from rat
bladder tissue. For a detailed description of the library
from which this clone was derived, please visit our web
site at ratest.eng.uiowa.edu. The subtraction has been
previously described in (Bonaldo, Lennon and Soares,
Genome Research 6:791-806, 1996)

KEYWORDS	GSS.
SOURCE	house mouse.
ORGANISM	Mus musculus
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus. 1 (bases 1 to 548)
AUTHORS	Dunn,D., Aoyagi,A., Barber,M., Beacorn,T., Duval,B., Hamil,C., Islam,H., Longacre,S., Mahmoud,M., Meenen,E., Pedersen,T., Reilly,M., Rose,M., Rose,R., Stokes,R., Tingey,A., von Niederhausern,A. and Wright,D., Weiss,R.
TITLE	Mouse whole genome scaffolding with paired end reads from 10kb plasmid inserts
JOURNAL	Unpublished (2000)
COMMENT	Contact: Robert B. Weiss University of Utah University of Utah Genome Center Rm. 308, Biomedical Polymers Research Bldg., 20 S. 2030 E., SLC, UT 84112, USA Tel: 801 585 5606 Fax: 801 585 7177 Email: ddunn@genetics.utah.edu Insert length: 10000 Std error: 0.00 Plate: 0226 row: M column: 20 Seq primer: CGTGTGAACGACGCGCCAGT Class: plasmid ends High quality sequence stop: 548. location/Qualifiers 1. 548 /organism="Mus musculus" /strain="C57BL/6J" /db_xref="taxon:10090" /clone="UUGC2M0226M20" /clone_1fb="Mouse 10kb plasmid UUGC2M library"
FEATURES	
source	

```

/sex="Female"
/ldb_host="E. coli strain XL10-Gold, Tl-resistant, F-"
/node="Vector: PWD42nv; Purified genomic DNA from M.
musculus C57BL/6J (female) was obtained from the Jackson
Laboratory Mouse DNA Resource
(http://www.jax.org/resources/documents/dnares/). The DNA
was hydrodynamically sheared by repeated passage through a
0.005 inch orifice at constant velocity. The sheared DNA
was blunt end-repaired with T4 DNA polymerase and T4
polynucleotide kinase. Adaptor oligonucleotides were
ligated to the blunt ends in high molar excess. The
adaptored DNA was purified and size-selected for a 9.5 to
10.5 kb range using preparative agarose gel
electrophoresis. Vector DNA was prepared from a derivative
of PWD42 (g14732114[gb|AF129072.1], a copy-number
inducible derivative of plasmid R1. The vector was ligated
with adaptors complementary to the insert adaptors and
purified. The sheared, adaptored mouse DNA was annealed to
adaptored vector DNA, and transformed into
chemically-competent E. coli XL10-Gold (Stratagene) cells
and selected for ampicillin resistance."

BASE COUNT      188 a      130 c      119 g      111 t
ORIGIN

Query Match      1.6%; Score 24; DB 12; Length 548;
Best Local Similarity 100.0%; Pred. No. 1.3e+03;
Matches 24; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1516 tctgcttacaataaaaaaa 1539
      |||||||
Db 93 TCTGCTTACAAAAA 116

RESULT 9
AK017944 1560 bp mRNA linear HTC 19-JAN-2002
LOCUS AK017944
DEFINITION Mus musculus adult male thymus cDNA, RIKEN full-length enriched
library, clone:5830424K06:hypothetical protein, full insert
sequence.

```



```

/lab_host="XLI-Blue MRF"/SOLR"
/Note="Vector: pBluescript; Site_1: EcoRI; Site_2: XhoI;
mRNA made from human bone marrow stroma. cDNA made by
oligo-dT priming. Directionally cloned. Size-selected for
average insert size >0.5 kb. Library supplied by Dr. Lidin
Jia (NHGR)."
BASE COUNT      50 a      36 c      39 g      73 t
ORIGIN

Query Match      1.5%; Score 23; DB 10; Length 198;
Best Local Similarity 100.0%; Pred. No. 4.9e+03;
Matches 23; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1517 ctgcttcaaaaaaaaaa 1539
Db 39 CTGCTTCAAAAAAAAAA 17

RESULT 11
LOCUS      A1535508      347 bp      mRNA      linear      EST 18-MAR-1999
DEFINITION UI-R-C3-sy-h-10-0-UI.s1 UI-R-C3 Rattus norvegicus cDNA clone
VERSION     A1535508
KEYWORDS    A1535508.1 GI:4449643
SOURCE      Norway rat.
ORGANISM    Rattus norvegicus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
            Rattus.
REFERENCE   1 (bases 1 to 347)
AUTHORS    Bonaldo,M.F., Lennon,G. and Soares,M.B.
TITLE      Normalization and subtraction: two approaches to facilitate gene
JOURNAL    Genome Res. 6 (9), 791-806 (1996)
MEDLINE    97044477
COMMENT     Contact: Soares, MB
            Program for Rat Gene Discovery and Mapping
            University of Iowa
            451 Eckstein Medical Research Building Iowa City, IA 52242, USA
            Tel: 319 335 8250
            Fax: 319 335 9565
            Email: mscores@blue.weeg.uiowa.edu
            The sequence contained an oligo-dT track that was present in the
            oligonucleotide that was used to prime the synthesis of first
            strand cDNA and therefore this may represent a bonafide poly A
            tail. cDNA library Preparation: M.B. Soares Lab Clone distribution:
            clones will be available through Research Genetics (www.resgen.com)
            Seq primer: M13 Forward
            Location/Qualifiers
                1..347
                /organism="Rattus norvegicus"
                /strain="Sprague-Dawley"
                /db_xref="taxon:10116"
                /clone="UI-R-C3-sy-h-10-0-UI"
                /clone_1lb="UI-R-C3"
                /dev_stage="adult"
                /lab_host="DH10B (Life Technologies)"
                /note="Vector: pT73D-Pac (Pharmacia) with a modified
                polylinker; Site_1: Not I; Site_2: Eco RI; The UI-R-C3
                library is a subtracted library of a series, ultimately
                derived from a mixture of individually tagged normalized
                libraries from rat placenta, adult lung, brain, liver,
                kidney, heart, spleen, ovary, muscle, and 8, 12 and 18-day
                embryos, after a series of subtractions to reduce the
                representation of cDNAs from which ESTs had already been
                generated. The following serially subtracted libraries
                were generated in this process: UI-R-C3, UI-R-C2p, UI-R-C1
                , UI-R-C0, UI-R-A1, UI-R-E1. The tag is a string of 3-5
                nucleotides present between the Not I site and the
                oligo-dT track which allows identification of the library

```

```

of origin of a clone within themixture. The subtracted
library (UI-R-C3) was constructed as follows: PCR amplified
cDNA inserts from UI-R-C2p clones from which 3' ESTs had
been derived was used as a driver in a hybridization with
the UI-R-C2p library in the form of single-stranded
circles. The remaining single-stranded circles (subtracted
library) was purified by hydroxyapatite column
chromatography, converted to double-stranded circles and
electroporated into DH10B bacteria (Life Technologies) to
generate the UI-R-C3 library. This procedure has been
previously described (Bonaldo, Lennon and Soares, Genome
Research 6:791-806, 1996)"
BASE COUNT      104 a      63 c      65 g      115 t
ORIGIN

Query Match      1.5%; Score 23; DB 9; Length 347;
Best Local Similarity 100.0%; Pred. No. 3.3e+03;
Matches 23; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1517 ctgcttcaaaaaaaaaa 1539
Db 23 CTGCTTCAAAAAAAAAA 1

RESULT 12
LOCUS      A1574766/c      351 bp      mRNA      linear      EST 30-MAR-1999
DEFINITION UI-R-G0-ub-e-05-0-UI.s1 UI-R-G0 Rattus norvegicus cDNA clone
VERSION     A1574766
KEYWORDS    A1574766.1 GI:4544762
SOURCE      Norway rat.
ORGANISM    Rattus norvegicus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
            Rattus.
REFERENCE   1 (bases 1 to 351)
AUTHORS    Bonaldo,M.F., Lennon,G. and Soares,M.B.
TITLE      Normalization and subtraction: two approaches to facilitate gene
JOURNAL    Genome Res. 6 (9), 791-806 (1996)
MEDLINE    97044477
COMMENT     Contact: Soares, MB
            Program for Rat Gene Discovery and Mapping
            University of Iowa
            451 Eckstein Medical Research Building Iowa City, IA 52242, USA
            Tel: 319 335 8250
            Fax: 319 335 9565
            Email: mscores@blue.weeg.uiowa.edu
            The sequence contained an oligo-dT track that was present in the
            oligonucleotide that was used to prime the synthesis of first
            strand cDNA and therefore this may represent a bonafide poly A
            tail. The sequence tag present in the cDNA between the NotI site
            and the oligo-dT track served to verify it as a clone from the
            normalized ganglia library cDNA library Preparation: M.B. Soares
            Lab Clone distribution: clones will be available through Research
            Genetics (www.resgen.com)
            Seq primer: M13 Forward
            Location/Qualifiers
                1..351
                /organism="Rattus norvegicus"
                /strain="Sprague-Dawley"
                /db_xref="taxon:10116"
                /clone="UI-R-G0-ub-e-05-0-UI"
                /clone_1lb="UI-R-G0"
                /dev_stage="adult"
                /lab_host="DH10B (Life Technologies)"
                /note="Vector: pT73D-Pac (Pharmacia) with a modified
                polylinker; Site_1: Not I; Site_2: Eco RI; The UI-R-G0
                library is a normalized library constructed from a
                mixture of rat tissues (nodose ganglia, dorsal root

```

ganglia, and trigeminal ganglia). The tag is a string of 6 nucleotides present between the Not I site and the oligo-dT track. The library was constructed as described by Bonaldo, Lennon and Soares, Genome Research 6: 791-806 '1996'

BASE COUNT 107 a 65 c 62 g 117 t

Query Match 1.5%; Score 23; DB 9; Length 351;
Best Local Similarity 100.0%; Pred. No. 3.3e+03;
Matches 23; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1517 ctgcttacaataaaataaa 1539

Db 23 CTGCTTACAAAAAATAAAAA 1

RESULT 13
LOCUS BF703043/c 359 bp mRNA linear EST 22-DEC-2000
DEFINITION MI-P-E5-abn-e-09-1-UM.s1 MI-P-E5 Sus scrofa cDNA clone
ACCESSION BF703043
VERSION BF703043.1 GI:11988451
KEYWORDS EST.
SOURCE pig.
ORGANISM Sus scrofa

REFERENCE 1 (bases 1 to 359)
AUTHORS Bonaldo,M.F., Lennon,G. and Soares,M.B.
TITLE Normalization and subtraction: two approaches to facilitate gene discovery
JOURNAL Genome Res. 6 (9), 791-806 (1996)
COMMENT 97044477
Contact: Tugale CK
Molecular Genetics Laboratory, Department of Animal Science
Iowa State University
201 Kildee Hall, Ames, IA 50011-3150, USA
Tel: 5152944252
Fax: 5152942401
Email: ctugale@iastate.edu

FEATURES
source

The sequence contained an oligo-dT track that was present in the oligonucleotide that was used to prime the synthesis of first strand cDNA and therefore this may represent a bonafide poly A tail. The sequence tag present in the cDNA between the NotI site and the oligo-dT track served to verify it as a clone from the non-normalized embryo at gestational day 12 library cDNA library preparation: RJ Woods, JA Green, RS Prather S142 Animal Science Research Center, Department of Animal Science, University of Missouri-Columbia, 65211 Clone distribution: clones will be available through Research Genetics (www.resgen.com)
Seq primer: M13 Forward
POLY-A=yes.

FEATURES
source

Location/Qualifiers
1. 359
/organism="Sus scrofa"
/strain="crossbred"
/db_xref="taxon:9823"
/clone="MI-P-E5-abn-e-09-1-UM"
/clone_1ib="MI-P-E5"
/lab_host="DH10B (Life Technologies)"
/note="Vector: pRT73D-Pac (Pharmacia) with a modified polylinker. Site:1: Not I; Site:2: EcoRI; The MI-P-E5 library is derived from embryo at gestational day 12. For a detailed description of the library from which this clone was derived, please visit our web site at
http://pigest.genome.iastate.edu/.

BASE COUNT 97 a 63 c 99 g 100 t

ORIGIN

Query Match 1.5%; Score 23; DB 10; Length 359;
Best Local Similarity 100.0%; Pred. No. 3.3e+03;
Matches 23; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1517 ctgcttacaataaaataaa 1539

Db 24 CTGCTTACAAAAAATAAAAA 2

RESULT 14
LOCUS D68664/c 360 bp mRNA linear EST 07-DEC-1995
DEFINITION CELK134B3F Yuji Kohara unpublished cDNA Caenorhabditis elegans CDNA
ACCESSION D68664
VERSION D68664.1 GI:1107386
KEYWORDS EST.
SOURCE Caenorhabditis elegans.
ORGANISM Caenorhabditis elegans.

REFERENCE 1 (bases 1 to 360)
AUTHORS Kohara,Y., Mitsuki,H., Nishigaki,A., Motohashi,T., Sugimoto,A. and Tabara,H.
TITLE Toward an expression map of the C.elegans genome
JOURNAL Unpublished (1994)
COMMENT Contact: Yuji Kohara
Genome Biology Lab.
National Institute of Genetics
Yata 1111, Mishima, Shizuoka 411, Japan
Tel: 81-559-81-6854
Fax: 81-559-81-6855
Email: ykohara@lab.nig.ac.jp.

FEATURES
source

Location/Qualifiers
1. 360
/organism="Caenorhabditis elegans"
/strain="CB1489 him-8(e1489)"
/db_xref="taxon:6239"
/clone="YK134B3"
/clone_1ib="Yuji Kohara unpublished cDNA"
/sex="hermaphrodite, male"
/tissue_type="whole animal"
/dev_stage="varied"

BASE COUNT 97 a 98 c 75 g 89 t 1 others

Query Match 1.5%; Score 23; DB 10; Length 360;
Best Local Similarity 100.0%; Pred. No. 3.2e+03;
Matches 23; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 337 tgcctgctgaatcccgatgag 359

Db 125 TGCTGGCTGAATCTCGATGAG 103

RESULT 15
LOCUS D63343/c 360 bp mRNA linear EST 07-DEC-1995
DEFINITION CELK068F8F Yuji Kohara unpublished cDNA Caenorhabditis elegans CDNA
ACCESSION D63343
VERSION D63343.1 GI:1104985
KEYWORDS EST.
SOURCE Caenorhabditis elegans.
ORGANISM Caenorhabditis elegans.

REFERENCE 1 (bases 1 to 360)
AUTHORS Kohara,Y., Mitsuki,H., Nishigaki,A., Motohashi,T., Sugimoto,A. and

TITLE Toward an expression map of the C.elegans genome
JOURNAL Unpublished (1994)
COMMENT Contact: Yuji Kohara
Genome Biology Lab.
National Institute of Genetics
Yata 1111, Mishima, Shizuoka 411, Japan
Tel: 81-559-81-6854
Fax: 81-559-81-6855
Email: ykohara@lab.nig.ac.jp.

FEATURES
Source
1. .360
/organism="Caenorhabditis elegans"
/strain="CB1489 him-8(e1489)"
/db_xref="taxon:6239"
/clone="yk68f8"
/clone_id="Yuji Kohara unpublished cDNA"
/sex="hermaphrodite, male"
/tissue_type="whole animal"
/dev_stage="varied"

BASE COUNT 94 a 95 c 74 g 89 t 8 others
ORIGIN

Query Match 1.5%; Score 23; DB 10; Length 360;
Best Local Similarity 100.0%; Pred. No. 3.2e+03;
Matches 23; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 337 Tgctgctgaatcgcgatgag 359
|||||
Db 133 TGCTGCTGAATCTCGATGAG 111

Search completed: July 18, 2002, 10:59:54
Job time: 2417 sec

83	20	1.3	31994	22	AAS30619	DNA encoding novel	156	19	1.2	535	22	AAI14987	Human breast cancer	
84	20	1.3	31994	22	AAS28165	Genomic sequence #	157	19	1.2	539	22	AAH68914	Human cervical can	
85	20	1.3	32035	22	AAK81068	Human immune/haema	158	19	1.2	543	22	AAH72554	Human cervical can	
86	20	1.3	57144	22	AAK79963	Human immune/haema	159	19	1.2	553	22	ABA21014	Human nervous syst	
c	87	20	1.3	113515	22	ABL34174	Human immune syste	160	19	1.2	553	22	ABA21015	Human nervous syst
c	88	19	1.2	21	16	AAO75644	Reverse transcript	161	19	1.2	555	22	ABA21018	Human nervous syst
c	89	19	1.2	24	22	AAI64873	Human serine/threo	162	19	1.2	555	21	AAE21646	Human breast and o
c	90	19	1.2	51	22	AAI28361	Human SNP oligonuc	163	19	1.2	555	22	AAK54916	Human haematologic
c	91	19	1.2	51	22	AAI28895	Human SNP oligonuc	164	19	1.2	555	22	AAE27239	Human secreted nove
c	92	19	1.2	51	22	AAI31945	Human SNP oligonuc	165	19	1.2	558	20	AAE27405	Human secreted pro
c	93	19	1.2	89	22	AAI98685	Human excretory re	166	19	1.2	568	22	ABA63030	Human foetal liver
c	94	19	1.2	94	22	AAI63081	Human kidney relat	167	19	1.2	568	22	ABA30330	Probe #8796 for ge
c	95	19	1.2	93	16	AAH88437	Cleavable padlock	168	19	1.2	568	22	AAK11489	Human brain expres
c	96	19	1.2	148	22	AAH7176	Human cervical can	169	19	1.2	568	22	AAK3727	Human bone marrow
c	97	19	1.2	185	22	AAK69821	Human immune/haema	170	19	1.2	568	22	AAI18106	Probe #8039 for ge
c	98	19	1.2	186	22	AAK57231	Human immune/haema	171	19	1.2	575	22	AAI43110	Probe #11796 used
c	99	19	1.2	234	22	AAK56917	Human immune/haema	172	19	1.2	575	22	AAH10397	Human cDNA clone (
c	100	19	1.2	254	22	AAH72083	Human cervical can	173	19	1.2	580	24	AAE96295	Arabidopsis cDNA e
c	101	19	1.2	265	22	AAI00613	Human reproductive	174	19	1.2	614	20	AAH81256	Human oligonuclein
c	102	19	1.2	266	22	ABK16193	Human lung tumour	175	19	1.2	621	22	AAH64775	Human secreted pro
c	103	19	1.2	275	21	AAA42928	Human secreted exp	176	19	1.2	624	21	AAK43906	Arabidopsis thalia
c	104	19	1.2	304	21	AAK19533	Human secreted pro	177	19	1.2	678	21	AAK59293	Human secreted pro
c	105	19	1.2	324	22	AAK34337	Human cDNA encodin	178	19	1.2	683	22	AAH98208	Rabbit EST-derived
c	106	19	1.2	338	19	AAV61226	CNA sequence of p	179	19	1.2	687	20	AAK30321	DNA encoding a hum
c	107	19	1.2	338	19	AAV58611	Prostate tumour sp	180	19	1.2	720	22	AAI08962	Human breast cancer
c	108	19	1.2	338	21	AAK06374	Human immunogenic	181	19	1.2	733	20	AAI25103	Human gene express
c	109	19	1.2	338	22	AAK63582	Human prostate CDN	182	19	1.2	735	21	AAA02248	Human colon cancer
c	110	19	1.2	338	22	AAK10133	Human prostate tum	183	19	1.2	760	20	AAI26256	Human gene express
c	111	19	1.2	338	22	AAH93490	Human prostate-spe	184	19	1.2	793	21	AAE52494	Human secreted pro
c	112	19	1.2	338	22	AAH84804	Human prostate-spe	185	19	1.2	805	21	AAK54186	Arabidopsis thalia
c	113	19	1.2	338	22	AAH02555	Human prostate-spe	186	19	1.2	821	22	AAH71485	Human cervical can
c	114	19	1.2	343	22	AAH08344	Human breast cancer	187	19	1.2	823	22	AAH19568	Human breast cancer
c	115	19	1.2	345	23	AAK57638	CDNA #314 encoding	188	19	1.2	863	21	AAK77640	Human cancer assoc
c	116	19	1.2	354	22	AAK79623	Human immune/haema	189	19	1.2	865	22	AAK67430	Human immune/haema
c	117	19	1.2	363	22	AAE65255	Novel human polynu	190	19	1.2	865	22	AAK67431	Human immune/haema
c	118	19	1.2	366	22	ABA47415	Human breast cell	191	19	1.2	865	22	AAK68170	Human immune/haema
c	119	19	1.2	366	22	ABA65303	Human foetal liver	192	19	1.2	865	22	AAK68172	Human immune/haema
c	120	19	1.2	366	22	ABA32404	Probe #10870 for g	193	19	1.2	871	22	AAI194007	Human neuroblastom
c	121	19	1.2	366	22	AAK13717	Human brain expres	194	19	1.2	878	19	AAV733447	Flea salivary protei
c	122	19	1.2	366	22	AAK39458	Human bone marrow	195	19	1.2	878	19	AAV733448	Flea salivary protei
c	123	19	1.2	366	22	AAI20268	Probe #10201 for g	196	19	1.2	890	22	AAI94042	Human neuroblastom
c	124	19	1.2	366	22	AAI45470	Probe #14156 used	197	19	1.2	899	22	AAE45108	Human secreted pro
c	125	19	1.2	366	22	AAI05972	Probe #5963 used t	198	19	1.2	905	22	AAH64812	Human secreted pro
c	126	19	1.2	368	22	AAI25580	Human breast cancer	199	19	1.2	912	22	AAI22666	Human breast cancer
c	127	19	1.2	369	22	AAI84575	Human polynucleoti	200	19	1.2	922	22	AAI16076	H. brasiliensis la
c	128	19	1.2	380	22	AAK57454	Human immune/haema	201	19	1.2	931	21	AAA59461	Nucleotide sequenc
c	129	19	1.2	383	22	AAI87502	Human polynucleoti	202	19	1.2	931	21	AAK53397	Clone seq. pK0037. h
c	130	19	1.2	393	22	AAE65925	Novel human polynu	203	19	1.2	931	21	AAE94953	Soybean glutathion
c	131	19	1.2	395	22	AAI86896	Human polynucleoti	204	19	1.2	931	21	AAE31576	Soybean type III G
c	132	19	1.2	397	22	AAI90049	Human polynucleoti	205	19	1.2	936	22	AAH70577	Human cervical can
c	133	19	1.2	403	22	AAI83197	Human polynucleoti	206	19	1.2	947	20	AAK86121	CDNA encoding a Tr
c	134	19	1.2	404	22	AAI86429	Human polynucleoti	207	19	1.2	960	20	AAA52713	Soybean Mlo homolo
c	135	19	1.2	407	22	AAI83508	Human polynucleoti	208	19	1.2	978	20	AAK51732	DNA encoding a hum
c	136	19	1.2	412	22	AAI86479	Human polynucleoti	209	19	1.2	998	22	ABAI1962	Human nervous syst
c	137	19	1.2	414	22	AAI86474	Human polynucleoti	210	19	1.2	1001	18	AAH787436	BCCP subunit of AC
c	138	19	1.2	415	22	AAI88881	Human polynucleoti	211	19	1.2	1011	20	AAKX00666	Human secreted pro
c	139	19	1.2	415	22	AAI90006	Human polynucleoti	212	19	1.2	1028	22	ABAI6563	Human nervous syst
c	140	19	1.2	415	22	AAI91424	Human polynucleoti	213	19	1.2	1029	22	ABAI6564	Human nervous syst
c	141	19	1.2	416	22	AAI81439	Human polynucleoti	214	19	1.2	1061	21	AAK77631	Human cancer assoc
c	142	19	1.2	417	22	AAK92617	Human cDNA 3'-end	215	19	1.2	1074	22	AAO05648	Human secreted pro
c	143	19	1.2	419	22	AAI81393	Human polynucleoti	216	19	1.2	1088	22	ABAI14674	Human nervous syst
c	144	19	1.2	437	22	AAH71012	Human cervical can	217	19	1.2	1107	22	AAE29972	Human lung antigen
c	145	19	1.2	440	22	AAI88046	Human polynucleoti	218	19	1.2	1107	22	AAE46222	Human cDNA encoding
c	146	19	1.2	448	22	AAI01742	Human reproductive	219	19	1.2	1116	21	AAAN02402	Cyclin dependent k
c	147	19	1.2	474	22	AAK35561	Human cardiovascular	220	19	1.2	1116	21	AAK95288	Soybean cyclin-dep
c	148	19	1.2	495	22	AAI23840	Human breast cancer	221	19	1.2	1119	21	AAK36809	DNA encoding a bra
c	149	19	1.2	501	22	AAK56459	Human cDNA for an	222	19	1.2	1133	21	AAA53791	Murine INGI 3' UTR
c	150	19	1.2	501	22	AAK68171	Human immune/haema	223	19	1.2	1134	20	AAK30368	DNA encoding a hum
c	151	19	1.2	515	22	AAI16737	Human breast cancer	224	19	1.2	1135	20	AAI219900	Human foetal kidn
c	152	19	1.2	529	22	AAK27548	CDNA encoding nove	225	19	1.2	1145	20	AAK61773	CDNA encoding huma
c	153	19	1.2	530	16	AAH45686	Medium chain-speci	226	19	1.2	1171	20	AAK61401	DNA encoding a hum
c	154	19	1.2	533	22	AAI09181	Human breast cancer	227	19	1.2	1178	21	AAK66439	Human secreted pro
c	155	19	1.2	534	22	AAK92396	Human cDNA 3'-end	228	19	1.2	1202	19	AAV69639	Translational init

229	19	1.2	1222	20	AAx83994	Human EPC1 encodin	302	19	1.2	3147	23	ABL02772	Drosophila melanog
230	19	1.2	1243	21	AAc95475	Human secreted pro	303	19	1.2	3153	22	AAK69887	Human immune/haema
231	19	1.2	1255	22	AAc08382	Human secreted pro	304	19	1.2	3163	22	AAK89450	Human digestive sy
232	19	1.2	1297	19	AAV73449	Fllea salivla protei	305	19	1.2	3164	22	AAK89447	Human digestive sy
233	19	1.2	1287	19	AAV73450	Fllea salivla protei	306	19	1.2	3146	22	ABL06440	Drosophila melanog
234	19	1.2	1302	19	AAV40485	Human secreted pro	307	19	1.2	3437	22	AAc92567	Mouse T2R26 nucleo
235	19	1.2	1307	21	AAc74248	Human secreted pro	308	19	1.2	3466	19	AAV37225	Human Bap-1 DNA.
236	19	1.2	1318	21	AAc55912	Runner bean gibber	309	19	1.2	3515	22	AAc60106	Arabidopsis thaila
237	19	1.2	1363	21	AAc55207	Human secreted pro	310	19	1.2	3577	23	AAc67293	DNA encoding novel
238	19	1.2	1376	21	AAc23399	Human secreted pro	311	19	1.2	3589	22	AAI60891	Human polynucleoti
239	19	1.2	1392	20	AAx04395	Human secreted pro	312	19	1.2	3774	21	AAc78005	Human cancer assoc
240	19	1.2	1417	21	AAc74398	Human secreted pro	313	19	1.2	3853	23	AAc77598	Human ORFX ORF3153
241	19	1.2	1516	22	AAc56392	ABC transport rela	314	19	1.2	3859	23	AAc76893	DNA encoding novel
242	19	1.2	1542	20	AAI16151	Human Sox1 encodin	315	19	1.2	3907	23	ABL26630	Drosophila melanog
243	19	1.2	1578	19	AAV71117	CDNA encoding a hu	316	19	1.2	4007	22	AAc67078	Human immune/haema
244	19	1.2	1578	19	AAV62388	Human 7-transmembr	317	19	1.2	4017	22	AAK72940	Human immune/haema
245	19	1.2	1584	21	AAc00015	Scorpion protein d	318	19	1.2	4058	22	AAc42016	Human small cell 1
246	19	1.2	1592	23	AAx07430	Homo sapiens secre	319	19	1.2	4091	22	AAc11111	Genomic sequence #
247	19	1.2	1599	23	AAc90606	DNA encoding novel	320	19	1.2	4189	23	AAc69085	Human encoding novel
248	19	1.2	1605	22	AAc94109	Human full-length	321	19	1.2	4196	20	AAc40488	Human SUD1 DNA.
249	19	1.2	1701	17	AAc30372	MORF-1 cDNA. Homo	322	19	1.2	4199	20	AAc34304	Human PRO286 nucle
250	19	1.2	1701	18	AAc61397	MORF-1 coding sequ	323	19	1.2	4199	20	AAc58296	Human TOLL protein
251	19	1.2	1701	21	AAc44745	Human FADD DNA. H	324	19	1.2	4199	21	AAc78584	Human PRO286 nucle
252	19	1.2	1736	21	AAc79724	Human secreted pro	325	19	1.2	4203	20	AAx07074	Human KDS1 serine/
253	19	1.2	1738	21	AAc59662	Human secreted pro	326	19	1.2	4335	23	ABL15141	Drosophila melanog
254	19	1.2	1821	21	AAc59421	Human secreted pro	327	19	1.2	4343	21	AAc77400	Human ORFX ORF2955
255	19	1.2	1835	21	AAc53789	Murine incl common	328	19	1.2	4354	23	ABL20674	Drosophila melanog
256	19	1.2	1857	21	AAc59142	Human secreted pro	329	19	1.2	4720	22	AAc24153	Human secreted pro
257	19	1.2	1860	21	AAc62425	Human brain-associ	330	19	1.2	4733	22	ABA07211	Human pancreatic c
258	19	1.2	1901	21	AAc64010	DNA encoding a hyd	331	19	1.2	4733	22	AAc89843	Human digestive sy
259	19	1.2	1901	21	AAc49780	Tomato hydropoxi	332	19	1.2	4814	23	ABL02008	Drosophila melanog
260	19	1.2	1916	21	AAc78064	Human cancer assoc	333	19	1.2	4874	20	AAc60103	Arabidopsis thaila
261	19	1.2	1916	22	AAH33274	Human colon cancer	334	19	1.2	5026	22	AAc46517	Tumour suppressor
262	19	1.2	1917	22	AAc05599	Human secreted pro	335	19	1.2	5044	23	AAc81832	DNA encoding novel
263	19	1.2	1933	10	AAc90120	Human lipocortin D	336	19	1.2	5234	24	ABL32945	Human immune syste
264	19	1.2	1943	22	AAc07717	Human secreted pro	337	19	1.2	5278	22	AAc46376	Tumour suppressor
265	19	1.2	1966	22	AAI59403	Human polynucleoti	338	19	1.2	5278	24	ABL32823	Human immune syste
266	19	1.2	1966	22	AAc07747	Human secreted pro	339	19	1.2	5328	24	ABL32560	Human immune syste
267	19	1.2	2001	18	AAc59700	PTH-like peptide D	340	19	1.2	5407	24	ABL34091	Human immune syste
268	19	1.2	2054	22	AAc03894	Human secreted pro	341	19	1.2	5413	22	AAc46694	Tumour suppressor
269	19	1.2	2060	22	AAc20949	Soybean arginyl-CK	342	19	1.2	5470	23	ABL03536	Drosophila melanog
270	19	1.2	2071	22	AAH33230	Human colon cancer	343	19	1.2	5518	24	ABL32216	Human immune syste
271	19	1.2	2192	23	ABL03537	Drosophila melanog	344	19	1.2	5546	24	ABL32390	Human immune syste
272	19	1.2	2215	22	AAH45148	Human protein kina	345	19	1.2	5688	22	AAc89950	Human bone marrow
273	19	1.2	2242	22	AAc94899	Human full-length	346	19	1.2	5763	24	ABL33288	Human immune syste
274	19	1.2	2289	24	AAc96307	Arabidopsis cDNA e	347	19	1.2	5922	24	ABL32450	Human immune syste
275	19	1.2	2288	24	AAc24790	Human secreted pro	348	19	1.2	5925	24	ABL33576	Human immune syste
276	19	1.2	2363	22	AAc74284	Common ice plant c	349	19	1.2	5931	22	AAc46703	Tumour suppressor
277	19	1.2	2376	21	AAc75837	Human ORFX ORF1392	350	19	1.2	5937	24	AAc61211	Human gene regulat
278	19	1.2	2476	22	AAH34123	Human colon cancer	351	19	1.2	6002	24	ABL32953	Human immune syste
279	19	1.2	2484	22	AAI59146	Human polynucleoti	352	19	1.2	6012	24	ABL34565	Human metastasis a
280	19	1.2	2507	21	AAc06561	Human immunogenic	353	19	1.2	6050	22	AAc46791	Tumour suppressor
281	19	1.2	2507	22	AAc63770	Human prostate cDN	354	19	1.2	6050	24	ABL34128	Human immune syste
282	19	1.2	2507	22	AAc93677	Human prostate-spe	355	19	1.2	6077	24	ABL33246	Human immune syste
283	19	1.2	2507	22	AAc84491	Prostate tumour an	356	19	1.2	6093	24	ABL34170	Human immune syste
284	19	1.2	2563	23	AAH02742	Arabidopsis cDNA e	357	19	1.2	6098	24	ABL32519	Human immune syste
285	19	1.2	2610	24	AAc96324	Arabidopsis cDNA e	358	19	1.2	6109	24	ABL32326	Human immune syste
286	19	1.2	2651	23	AAc87521	Arabidopsis cDNA e	359	19	1.2	6109	24	AAc61077	Human immune syste
287	19	1.2	2651	23	AAc64607	DNA encoding novel	360	19	1.2	6160	24	ABL33533	Human gene regulat
288	19	1.2	2652	23	AAc23292	Drosophila melanog	361	19	1.2	6174	24	ABL33509	Human immune syste
289	19	1.2	2653	23	AAc69572	Human immune/haema	362	19	1.2	6189	22	AAc46599	Human immune syste
290	19	1.2	2663	23	ABL18654	Drosophila melanog	363	19	1.2	6189	24	ABL34592	Human immune syste
291	19	1.2	2712	22	AAH75376	P. hybida pollen-	364	19	1.2	6270	22	AAc45307	Human metastasis a
292	19	1.2	2712	22	AAH45742	Petunia zinc fling	365	19	1.2	6270	24	ABL32405	Chemically pretrea
293	19	1.2	2738	22	AAc06007	Human neuronal apo	366	19	1.2	6283	24	AAc61351	Human immune syste
294	19	1.2	2743	24	AAc966335	Arabidopsis cDNA e	367	19	1.2	6412	24	AAc61145	Human gene regulat
295	19	1.2	2817	21	AAc53790	Murine P37ING1 cod	368	19	1.2	6465	24	ABL32985	Human gene regulat
296	19	1.2	2962	20	AAc82161	Rat Acid sensitive	369	19	1.2	6558	22	AAc46549	Human immune syste
297	19	1.2	3037	23	ABL20562	Drosophila melanog	370	19	1.2	6577	22	AAc46719	Tumour suppressor
298	19	1.2	3038	23	ABL03730	Drosophila melanog	371	19	1.2	6577	22	AAc33357	Tumour suppressor
299	19	1.2	3048	22	AAc69888	Human immune/haema	372	19	1.2	6577	24	AAc61222	Human gene regulat
300	19	1.2	3050	22	AAc69889	Human immune/haema	373	19	1.2	6591	24	ABL33251	Human immune syste
301	19	1.2	3101	11	AAc02047	Sequence encoding	374	19	1.2	6627	24	ABL33575	Human immune syste

C 375	19	1.2	6823	24	ABL33145	Human immune syste	C 448	19	1.2	18817	24	ABL34495	Human metastasis a
C 376	19	1.2	6853	24	ABL33211	Human immune syste	C 449	19	1.2	19082	24	ABL32626	Human immune syste
C 377	19	1.2	6915	24	ABL32406	Human immune syste	C 450	19	1.2	19131	24	AA546716	Tumour suppressor
C 378	19	1.2	7010	22	ABL03343	Human reproductive	C 451	19	1.2	19972	24	ABL32106	Human immune syste
C 379	19	1.2	7273	23	ABL06708	Drosophila melanog	C 452	19	1.2	21446	22	AAK70163	Human immune/haema
C 380	19	1.2	7241	24	AA561395	Human gene regulat	C 453	19	1.2	21446	22	AAK82629	Human immune/haema
C 381	19	1.2	7449	24	ABL32277	Human immune syste	C 454	19	1.2	22408	23	ABL04054	Drosophila melanog
C 382	19	1.2	7657	22	AA545477	Chemically pretrea	C 455	19	1.2	22874	23	ABL04054	Drosophila melanog
C 383	19	1.2	7657	22	AA545478	Chemically pretrea	C 456	19	1.2	22967	23	ABL12942	Drosophila melanog
C 384	19	1.2	7657	24	ABL34022	Human immune syste	C 457	19	1.2	23203	23	ABL17834	Drosophila melanog
C 385	19	1.2	7812	24	ABL34023	Human immune syste	C 458	19	1.2	23683	24	ABL34623	Human metastasis a
C 386	19	1.2	7857	22	ABL34023	Human immune syste	C 459	19	1.2	23899	23	ABL10362	Drosophila melanog
C 387	19	1.2	8030	24	ABL33333	Genomic sequence #	C 460	19	1.2	37973	24	ABL34196	Human immune syste
C 388	19	1.2	8078	22	AA546492	Human immune syste	C 461	19	1.2	39353	22	AAK65396	Human immune/haema
C 389	19	1.2	8136	22	AA531510	Tumour suppressor	C 462	19	1.2	39353	22	AAK67232	Human immune/haema
C 390	19	1.2	8197	24	ABL34515	Human DNA for a no	C 463	19	1.2	39358	22	AAK65397	Human immune/haema
C 391	19	1.2	8246	24	ABL33202	Human metastasis a	C 464	19	1.2	39358	22	AAK67233	Human immune/haema
C 392	19	1.2	8277	24	ABL33328	Human immune syste	C 465	19	1.2	43795	21	AA292583	Human immune/haema
C 393	19	1.2	8297	24	ABL34074	Human immune syste	C 466	19	1.2	48037	22	AAK84729	Human immune/haema
C 394	19	1.2	8499	23	ABL16018	Human immune syste	C 467	19	1.2	48037	22	AAK84729	Human immune/haema
C 395	19	1.2	8513	22	AA545355	Drosophila melanog	C 468	19	1.2	49380	23	ABL11838	Drosophila melanog
C 396	19	1.2	8513	22	AA545355	Chemically pretrea	C 469	19	1.2	49380	23	ABL11838	Drosophila melanog
C 397	19	1.2	8676	24	AA561414	Tumour suppressor	C 470	19	1.2	52691	22	AAK84435	Human immune/haema
C 398	19	1.2	8771	24	ABL33824	Human gene regulat	C 471	19	1.2	68940	20	AAK57351	Human immune/haema
C 399	19	1.2	8771	24	ABL33825	Human immune syste	C 472	19	1.2	110000	22	AAK84800	Human chromosome 6
C 400	19	1.2	8876	24	ABL34076	Human immune syste	C 473	19	1.2	113515	24	ABL34175	Nucleotide sequenc
C 401	19	1.2	8991	24	AA561295	Human immune syste	C 474	19	1.2	144460	21	AA293815	Human immune syste
C 402	19	1.2	9122	22	AA561295	Human gene regulat	C 475	19	1.2	465237	24	ABA90193	Olfactory receptor
C 403	19	1.2	9219	22	AA546807	Human musculoskele	C 476	19	1.2	1503900	22	AAK95240	Human oestrogen re
C 404	19	1.2	9219	22	AA546807	Tumour suppressor	C 477	19	1.2	1503900	22	AAK95240	Human neurigin-1
C 405	19	1.2	9219	24	ABL34426	Human immune syste	C 478	19	1.2	20	16	AAK6773	Human neurigin-1
C 406	19	1.2	9375	22	AAK37084	Human musculoskele	C 479	18	1.2	20	16	AAK75568	Reverse transcript
C 407	19	1.2	9375	22	AAK84379	Human immune/haema	C 480	18	1.2	21	16	AAK83959	BAV28 gene fragmen
C 408	19	1.2	9502	22	AA546731	Tumour suppressor	C 481	18	1.2	21	16	AAO75643	Reverse transcript
C 409	19	1.2	9577	23	ABL23106	Drosophila melanog	C 482	18	1.2	21	16	AAO75643	Reverse transcript
C 410	19	1.2	9741	24	ABL33323	Human immune syste	C 483	18	1.2	21	16	AAO75646	Reverse transcript
C 411	19	1.2	10034	24	ABL33885	Human immune syste	C 484	18	1.2	24	24	AB185028	Capture oligonucle
C 412	19	1.2	10189	24	ABL34161	Human immune syste	C 485	18	1.2	24	24	AB185029	Capture oligonucle
C 413	19	1.2	10191	24	ABL34255	Human immune syste	C 486	18	1.2	25	21	AAK85029	HLA DPA1 gene PCR
C 414	19	1.2	10647	24	AA561396	Human gene regulat	C 487	18	1.2	29	18	AAK69832	Probe for tandem s
C 415	19	1.2	10717	24	ABL33694	Human immune syste	C 488	18	1.2	30	17	AAK24223	Oligonucleotide pr
C 416	19	1.2	11131	24	ABL32753	Human immune syste	C 489	18	1.2	30	17	AAK24222	Oligonucleotide pr
C 417	19	1.2	11398	24	ABL34438	Human immune syste	C 490	18	1.2	30	17	AAK24224	Oligonucleotide pr
C 418	19	1.2	11836	22	AA545394	Chemically pretrea	C 491	18	1.2	50	22	AAK31421	Human SNP oligonuc
C 419	19	1.2	11996	24	ABL34493	Human metastasis a	C 492	18	1.2	50	22	AAK31422	Human SNP oligonuc
C 420	19	1.2	12138	24	ABL33943	Human immune syste	C 493	18	1.2	50	22	AAK31423	Human SNP oligonuc
C 421	19	1.2	12248	23	ABL14532	Drosophila melanog	C 494	18	1.2	51	22	AAK31428	Human SNP oligonuc
C 422	19	1.2	12492	21	AA550353	Human Goodpasture	C 495	18	1.2	51	22	AAK31661	Human SNP oligonuc
C 423	19	1.2	12730	19	AAV44233	Drosophila melanog	C 496	18	1.2	51	22	AAK32366	Human SNP oligonuc
C 424	19	1.2	12806	23	ABL24266	Lettuce resistance	C 497	18	1.2	51	22	AAK32398	Human SNP oligonuc
C 425	19	1.2	13214	22	AAK18377	Drosophila melanog	C 498	18	1.2	52	21	AAK15328	Human secreted pro
C 426	19	1.2	13333	22	AAK79541	Human nervous syst	C 499	18	1.2	70	17	AAK17035	Human mitochondria
C 427	19	1.2	13712	24	ABL33531	Human immune/haema	C 500	18	1.2	71	21	AAK29601	Human secreted pro
C 428	19	1.2	14169	22	AAK72936	Human immune syste	C 501	18	1.2	72	22	AAK80775	Recombinant RNA ca
C 429	19	1.2	14307	24	ABL32728	Human immune/haema	C 502	18	1.2	72	22	AAK66829	Human immune/haema
C 430	19	1.2	14686	23	ABL15140	Human immune syste	C 503	18	1.2	72	22	AAK80857	Human immune/haema
C 431	19	1.2	14925	22	AAK03282	Drosophila melanog	C 504	18	1.2	72	22	AAK83532	Human immune/haema
C 432	19	1.2	14925	22	AAK04552	Human reproductive	C 505	18	1.2	78	21	AAK86588	Human immune/haema
C 433	19	1.2	15051	23	ABL03732	Human reproductiv	C 506	18	1.2	88	22	AAK14386	Human secreted pro
C 434	19	1.2	15306	22	AA545454	Drosophila melanog	C 507	18	1.2	88	22	AAK198803	Human excretory re
C 435	19	1.2	15548	22	ABL34155	Chemically pretrea	C 508	18	1.2	94	18	AAV00411	Human kidney relat
C 436	19	1.2	15633	24	ABL03060	Human immune syste	C 509	18	1.2	95	22	AAK04913	3' fragment of C10
C 437	19	1.2	16217	24	ABL34157	Drosophila melanog	C 510	18	1.2	112	16	AAK25481	Human gene signatu
C 438	19	1.2	16766	24	ABL34157	Human immune syste	C 511	18	1.2	117	21	AAK16447	Human prostate can
C 439	19	1.2	16842	22	AA546411	Human immune syste	C 512	18	1.2	117	22	AAK68873	Human immune/haema
C 440	19	1.2	16842	24	AA561335	Tumour suppressor	C 513	18	1.2	117	22	AAK76691	Human immune/haema
C 441	19	1.2	16891	24	ABL32653	Human gene regulat	C 514	18	1.2	117	22	AAK83853	Human immune/haema
C 442	19	1.2	16891	24	ABL34505	Human immune syste	C 515	18	1.2	125	22	AAK81138	Human immune/haema
C 443	19	1.2	17137	24	ABL332191	Human metastasis a	C 516	18	1.2	137	22	AAK15347	Human immune/haema
C 444	19	1.2	17389	24	ABL33415	Human immune syste	C 517	18	1.2	137	22	AAK15347	Human breast cance
C 445	19	1.2	17738	24	ABL33538	Human immune syste	C 518	18	1.2	139	22	AAK17705	Human nervous syst
C 446	19	1.2	18296	23	ABL03234	Drosophila melanog	C 519	18	1.2	139	22	AAK17705	Human nervous syst
C 447	19	1.2	18512	24	ABL32976	Human immune syste	C 520	18	1.2	141	22	AAK15776	Human musculoskele

C 521	18	1.2	141	22	AAK83995	Human immune/haema	C 594	18	1.2	320	22	AAU13401	Human breast cancer
C 522	18	1.2	141	22	AAK87112	Human immune/haema	C 595	18	1.2	321	24	AA561485	Lung small cell ca
C 523	18	1.2	142	22	AAU15315	Human breast cancer	C 596	18	1.2	328	22	AA186509	Human polynucleoti
C 524	18	1.2	142	22	AAU24161	Human breast cancer	C 597	18	1.2	331	22	AAK54664	Human haematologic
C 525	18	1.2	142	22	AAK70077	Human immune/haema	C 598	18	1.2	331	22	AAK64191	Human immune/haema
C 526	18	1.2	143	22	AAK65842	Human immune/haema	C 599	18	1.2	333	22	AAU14965	Human breast cancer
C 527	18	1.2	143	22	AAK90707	Human digestive sy	C 600	18	1.2	336	22	AAK69254	Human immune/haema
C 528	18	1.2	146	22	AAK87315	Human immune/haema	C 601	18	1.2	338	22	AAK35462	Human cardiovascular
C 529	18	1.2	148	22	AAK32635	Human genomic DNA	C 602	18	1.2	339	22	AAU35334	Human musculoskele
C 530	18	1.2	159	22	AAK75654	Human immune/haema	C 603	18	1.2	342	22	AA187251	Human polynucleoti
C 531	18	1.2	160	18	AAU61985	Detection probe fo	C 604	18	1.2	342	22	AA188836	Human polynucleoti
C 532	18	1.2	160	22	AAU40236	DNA encoding human	C 605	18	1.2	344	22	AA187135	Human polynucleoti
C 533	18	1.2	160	22	AAU02154	Human reproductive	C 606	18	1.2	344	24	AA596303	Arabidopsis cDNA e
C 534	18	1.2	164	22	AA199160	Human excretory re	C 607	18	1.2	349	22	AA180809	Human polynucleoti
C 535	18	1.2	164	22	AA163510	Human kidney relat	C 608	18	1.2	354	21	AAK63882	Human polynucleoti
C 536	18	1.2	165	21	AAU14177	Human secreted pro	C 609	18	1.2	356	22	ABA17877	Nucleotide sequenc
C 537	18	1.2	165	22	ABA15465	Human nervous syst	C 610	18	1.2	359	22	AA188709	Human nervous syst
C 538	18	1.2	165	22	AAK64234	Human immune/haema	C 611	18	1.2	361	22	AA184186	Human polynucleoti
C 539	18	1.2	170	20	AAH85961	Human single nucle	C 612	18	1.2	363	22	AA532910	Human genomic DNA
C 540	18	1.2	172	22	AAK81214	Human immune/haema	C 613	18	1.2	367	22	AAU18612	Human breast cancer
C 541	18	1.2	177	22	AAK65967	Human immune/haema	C 614	18	1.2	369	22	AA190989	Human polynucleoti
C 542	18	1.2	180	21	AAU55658	Human differential	C 615	18	1.2	371	14	AAO61288	Human polynucleoti
C 543	18	1.2	181	20	AAU13491	Human gene express	C 616	18	1.2	371	23	AA585359	Human brain Expres
C 544	18	1.2	192	21	AAU12149	Human secreted pro	C 617	18	1.2	375	22	AA187531	Human polynucleoti
C 545	18	1.2	192	22	AAK77059	Human immune/haema	C 618	18	1.2	375	22	AA190980	Human polynucleoti
C 546	18	1.2	192	22	AAK77060	Human immune/haema	C 619	18	1.2	378	22	AA181567	Human polynucleoti
C 547	18	1.2	211	22	AAK67607	Human immune/haema	C 620	18	1.2	380	22	AA182817	Human polynucleoti
C 548	18	1.2	217	22	AAK82434	Human ovarian PCR-	C 621	18	1.2	380	22	AA186586	Human polynucleoti
C 549	18	1.2	218	22	AAK824704	Human ovarian tumor	C 622	18	1.2	380	22	AAK57792	Human immune/haema
C 550	18	1.2	218	22	AAK83332	Human ovarian tumor	C 623	18	1.2	381	22	AA188668	Human polynucleoti
C 551	18	1.2	220	22	AAK65830	Human immune/haema	C 624	18	1.2	382	22	AA187149	Human polynucleoti
C 552	18	1.2	220	22	AAK65820	Human immune/haema	C 625	18	1.2	382	22	AA187691	Human polynucleoti
C 553	18	1.2	224	19	AAU05742	Nucleotide sequenc	C 626	18	1.2	385	22	AA193254	Human polynucleoti
C 554	18	1.2	224	22	AAU24213	Human ovarian PCR-	C 627	18	1.2	387	21	AAZ94886	Rat taste specific
C 555	18	1.2	224	22	AAH82794	Human ovarian tumor	C 628	18	1.2	389	22	AA187865	Human polynucleoti
C 556	18	1.2	226	22	AAU34674	Human musculoskele	C 629	18	1.2	389	22	AA188391	Human polynucleoti
C 557	18	1.2	231	21	AAZ80203	Human colon cancer	C 630	18	1.2	389	22	AAK55411	Human immune/haema
C 558	18	1.2	232	22	AAH87963	Peppermint plant o	C 631	18	1.2	390	22	AAK54111	Human polynucleoti
C 559	18	1.2	240	22	AAU24341	Human ovarian PCR-	C 632	18	1.2	390	22	AAU09940	Human breast cancer
C 560	18	1.2	241	22	AAH82930	Human ovarian tumor	C 633	18	1.2	391	22	AAU17484	Human polynucleoti
C 561	18	1.2	242	22	AAU24979	Human ovarian PCR-	C 634	18	1.2	391	22	AA183168	Human polynucleoti
C 562	18	1.2	242	22	AAH83624	Human ovarian tumor	C 635	18	1.2	392	22	AA186653	Human polynucleoti
C 563	18	1.2	247	21	AAU93738	Cat flea hindgut a	C 636	18	1.2	395	22	AAH35297	Human polynucleoti
C 564	18	1.2	247	21	AAU94506	Cat flea hindgut a	C 637	18	1.2	397	22	AA185922	Human polynucleoti
C 565	18	1.2	249	22	AAU10931	Human breast cancer	C 638	18	1.2	398	22	AA183988	Human polynucleoti
C 566	18	1.2	249	22	AAU18695	Human breast cancer	C 639	18	1.2	398	22	AAK73447	Human immune/haema
C 567	18	1.2	249	22	AAK79399	Human immune/haema	C 640	18	1.2	399	22	AA185422	Human polynucleoti
C 568	18	1.2	251	22	AAU10961	Human breast cancer	C 641	18	1.2	399	22	AA187654	Human polynucleoti
C 569	18	1.2	251	22	AAU198673	Human excretory re	C 642	18	1.2	400	21	AAK93683	Cat flea hindgut a
C 570	18	1.2	253	22	AAU27556	CDNA encoding nove	C 643	18	1.2	401	22	AAK96165	Human neuregulin g
C 571	18	1.2	253	22	AAU163069	Human kidney relat	C 644	18	1.2	401	22	AAK96166	Human neuregulin g
C 572	18	1.2	258	20	AAU09018	Differentiation in	C 645	18	1.2	401	22	AAK96200	Human neuregulin g
C 573	18	1.2	258	22	AAK65732	Human immune/haema	C 646	18	1.2	401	22	AAK96708	Human neuregulin g
C 574	18	1.2	270	22	AAK67467	Human immune/haema	C 647	18	1.2	401	22	AAK97658	Human neuregulin g
C 575	18	1.2	274	21	AAU00718	Human colon cancer	C 648	18	1.2	401	22	AAK97659	Human neuregulin g
C 576	18	1.2	276	20	AAU89725	EST clone CT706.	C 649	18	1.2	401	22	AAK97693	Human neuregulin g
C 577	18	1.2	279	22	AAU24357	Human ovarian PCR-	C 650	18	1.2	401	22	AAK98201	Human neuregulin g
C 578	18	1.2	279	22	AAH82947	Human ovarian tumor	C 651	18	1.2	403	22	AAU192107	Human polynucleoti
C 579	18	1.2	281	22	AAU28986	Colon tumour relat	C 652	18	1.2	403	22	AAK96719	Human neuregulin g
C 580	18	1.2	284	22	AAU29083	CDNA encoding for	C 653	18	1.2	403	22	AAK98212	Human neuregulin g
C 581	18	1.2	290	22	AAU11318	Human breast cancer	C 654	18	1.2	403	22	AAU12000	Human breast cancer
C 582	18	1.2	294	22	AAK66970	Human immune/haema	C 655	18	1.2	403	22	AAU182265	Human polynucleoti
C 583	18	1.2	295	22	AAU23818	Human breast cancer	C 656	18	1.2	404	22	AAU09899	Human breast cancer
C 584	18	1.2	296	14	AAU059461	Human brain Expres	C 657	18	1.2	405	22	AA188744	Human polynucleoti
C 585	18	1.2	300	22	AAU56393	ABC transport rela	C 658	18	1.2	405	22	AAK96662	Human neuregulin g
C 586	18	1.2	302	22	AAK63052	Human immune/haema	C 659	18	1.2	407	22	AAK98135	Human neuregulin g
C 587	18	1.2	304	14	AAU05952	Human brain Expres	C 660	18	1.2	407	22	AAU15488	Human breast cancer
C 588	18	1.2	304	22	AAU05753	Human reproductive	C 661	18	1.2	409	22	AA189271	Human polynucleoti
C 589	18	1.2	304	22	AAK84486	Human immune/haema	C 662	18	1.2	412	22	AA185069	Human polynucleoti
C 590	18	1.2	305	22	AAK63278	Human immune/haema	C 663	18	1.2	412	22	AA186509	Human polynucleoti
C 591	18	1.2	306	22	AAU20217	Human breast cancer	C 664	18	1.2	413	22	AAK65941	Human immune/haema
C 592	18	1.2	308	22	AAK70443	Human immune/haema	C 665	18	1.2	415	22	AA188281	Human polynucleoti
C 593	18	1.2	314	22	AAU00786	Human reproductive	C 666	18	1.2	419	21	AAU59229	Human secreted pro

667	18	1.2	419	21	AAC69548	Human secreted pro
668	18	1.2	419	22	AA180417	Human polynucleoti
669	18	1.2	419	22	AAS02417	Human secreted pro
670	18	1.2	420	22	AA101117	Human breast cancer
671	18	1.2	421	21	AA280247	Human colon cancer
672	18	1.2	423	22	AA185265	Human polynucleoti
673	18	1.2	425	21	AA437895	EST sequence #7.
674	18	1.2	425	22	AA186980	Human polynucleoti
675	18	1.2	426	24	AAS96331	Arabidopsis CDNA e
676	18	1.2	427	22	AA193379	Human polynucleoti
677	18	1.2	431	22	AA187912	Human polynucleoti
678	18	1.2	434	22	AA124333	Human breast cancer
679	18	1.2	437	22	AAK87407	Human immune/haema
680	18	1.2	438	20	AAZ06770	PDE homologous seq
681	18	1.2	439	22	AAK55923	Human immune/haema
682	18	1.2	442	22	AA185193	Human polynucleoti
683	18	1.2	443	22	ABA57053	Human foetal liver
684	18	1.2	443	22	AAK88750	Human digestive sy
685	18	1.2	443	22	AA531784	Human liver associ
686	18	1.2	446	22	AA137502	Human musculoskele
687	18	1.2	446	22	AA137503	Human musculoskele
688	18	1.2	448	22	AA122802	Human breast cancer
689	18	1.2	450	22	AA117849	Human breast cancer
690	18	1.2	452	22	AA102233	Human reproductive
691	18	1.2	452	22	AA162531	Human breast or ov
692	18	1.2	456	20	AAV86590	EST clone BG3. Ho
693	18	1.2	457	22	AA198810	Human excretory re
694	18	1.2	457	22	AA135597	Human musculoskele
695	18	1.2	457	22	AA163206	Human kidney relat
696	18	1.2	460	22	AAH34388	Human colon cancer
697	18	1.2	463	22	AA192884	Human polynucleoti
698	18	1.2	464	22	AA186425	Human polynucleoti
699	18	1.2	464	22	AA193001	Human polynucleoti
700	18	1.2	470	22	AAK66310	Human immune/haema
701	18	1.2	470	22	AAK66311	Human immune/haema
702	18	1.2	474	19	AAV59513	Human secreted pro
703	18	1.2	476	22	AA136728	Human musculoskele
704	18	1.2	477	22	ABA59173	Human foetal liver
705	18	1.2	477	22	ABA27943	Probe #6409 for ge
706	18	1.2	477	22	AAK07364	Human brain expres
707	18	1.2	477	22	AAK33131	Human bone marrow
708	18	1.2	477	22	AA116345	Probe #6278 for ge
709	18	1.2	477	22	AA138927	Probe #7613 used t
710	18	1.2	479	22	ABA42189	Human breast cell
711	18	1.2	479	22	ABA52611	Human foetal liver
712	18	1.2	479	22	ABA22399	Probe #865 for gen
713	18	1.2	479	22	AAK00872	Human brain expres
714	18	1.2	479	22	AAK26327	Human bone marrow
715	18	1.2	479	22	AA110860	Probe #893 for gen
716	18	1.2	479	22	AA132220	Probe #906 used to
717	18	1.2	479	22	AA100881	Probe #872 used to
718	18	1.2	481	22	AA116003	Human breast cancer
719	18	1.2	486	22	AA196933	Human Mammary Glan
720	18	1.2	489	22	AA113840	Human breast cancer
721	18	1.2	492	18	AA184952	Human prostate pro
722	18	1.2	492	20	AAK35882	CDNA encoding a pr
723	18	1.2	495	20	AA224821	Human secreted pro
724	18	1.2	496	22	AA114955	Human breast cancer
725	18	1.2	496	22	AA112047	Human breast cancer
726	18	1.2	499	21	AAK59372	Human secreted pro
727	18	1.2	501	22	AAK62927	Human immune/haema
728	18	1.2	503	22	AA113107	Human breast cancer
729	18	1.2	507	22	AA535593	Human cardiovascular
730	18	1.2	508	22	ABA14314	Human nervous syst
731	18	1.2	508	22	ABA07689	Human ovarian and
732	18	1.2	508	22	AA102641	Human reproductive
733	18	1.2	515	22	AAH09428	Human CDNA clone (
734	18	1.2	516	22	ABA62628	Human foetal liver
735	18	1.2	516	22	ABA29942	Probe #8408 for ge
736	18	1.2	516	22	AAK10996	Human brain expres
737	18	1.2	516	22	AAK36838	Human bone marrow
738	18	1.2	516	22	AA117689	Probe #7622 for ge
739	18	1.2	516	22	AA142623	Probe #11309 used
740	18	1.2	517	22	AA180518	Human polynucleoti
741	18	1.2	519	20	AAK61880	Human immune/haema
742	18	1.2	521	20	AAV86693	EST clone B166. H
743	18	1.2	521	24	AA561743	Lung small cell ca
744	18	1.2	522	22	AA007907	Human secreted pro
745	18	1.2	525	22	AAK61793	Human immune/haema
746	18	1.2	526	16	AA086309	Wilson disease gen
747	18	1.2	527	22	AAK73020	Human immune/haema
748	18	1.2	529	22	AAK73018	Human immune/haema
749	18	1.2	530	22	AA556473	Human CDNA for an
750	18	1.2	530	22	AAK73019	Human immune/haema
751	18	1.2	540	22	AA105347	Human reproductive
752	18	1.2	540	22	AAK71626	Human immune/haema
753	18	1.2	544	22	AAH10053	Human CDNA clone (
754	18	1.2	546	22	ABA60636	Human foetal liver
755	18	1.2	546	22	ABA28743	Probe #7209 for ge
756	18	1.2	546	22	AAK08918	Human brain expres
757	18	1.2	546	22	AAK34808	Human bone marrow
758	18	1.2	546	22	AA116835	Probe #6768 for ge
759	18	1.2	546	22	AA140527	Probe #9213 used t
760	18	1.2	548	22	AAH13288	Human CDNA clone (
761	18	1.2	548	22	AAK58546	Human immune/haema
762	18	1.2	552	22	AAK33132	Human secreted pro
763	18	1.2	555	22	AAH09478	Human CDNA clone (
764	18	1.2	555	22	AAH09942	Human CDNA clone (
765	18	1.2	560	22	AA102430	Human reproductive
766	18	1.2	560	22	AAK64677	Human immune/haema
767	18	1.2	564	22	AAH10298	Human CDNA clone (
768	18	1.2	567	22	AAK62114	Human immune/haema
769	18	1.2	567	22	AAK76252	Human immune/haema
770	18	1.2	567	22	AAK76253	Human immune/haema
771	18	1.2	568	22	AAK81769	Human immune/haema
772	18	1.2	569	22	AAH10316	Human CDNA clone (
773	18	1.2	570	22	AAH33656	Human colon cancer
774	18	1.2	571	21	AAK22330	Human secreted pro
775	18	1.2	573	22	AAK80470	Human immune/haema
776	18	1.2	575	22	AAK27483	CDNA encoding nove
777	18	1.2	582	21	AAH11948	Human cytokine rec
778	18	1.2	583	22	AAH10021	Human CDNA clone (
779	18	1.2	586	22	AAK71625	Human immune/haema
780	18	1.2	594	20	AAV88419	EST clone EK634.
781	18	1.2	599	22	AAK34743	Human DNA for a no
782	18	1.2	599	20	AAZ06774	Human phosphodiester
783	18	1.2	599	23	AA577758	DNA encoding novel
784	18	1.2	600	20	AAK30350	DNA encoding a hum
785	18	1.2	600	22	AA120932	Human breast cancer
786	18	1.2	600	22	AAH64831	Human secreted pro
787	18	1.2	604	22	AAK30544	DNA encoding novel
788	18	1.2	604	22	AA106320	Human reproductive
789	18	1.2	608	20	AAK20439	Human secreted pro
790	18	1.2	611	11	AAQ03848	Equine beta LH sub
791	18	1.2	612	22	AAK53246	Human CDNA encodin
792	18	1.2	618	22	AAK81393	Human immune/haema
793	18	1.2	619	21	AAK50674	Hermansky-Pudlak s
794	18	1.2	621	22	AAH34680	Human colon cancer
795	18	1.2	624	21	AAH16409	Human colon cancer
796	18	1.2	629	22	AAK75088	Human colon associ
797	18	1.2	633	22	ABA19495	Human nervous syst
798	18	1.2	634	22	AAK80474	Human immune/haema
799	18	1.2	635	22	AAK80472	Human immune/haema
800	18	1.2	636	21	AAK59368	Human secreted pro
801	18	1.2	641	23	AAK591349	DNA encoding novel
802	18	1.2	642	20	AAK75040	Human secreted pro
803	18	1.2	646	21	AAH81659	N. meningitidis pa
804	18	1.2	649	20	AAK24522	Human lung tumor a
805	18	1.2	649	21	AAK65761	Human lung cancer
806	18	1.2	650	22	AAK64518	Human immune/haema
807	18	1.2	652	22	AAH15611	Human nervous syst
808	18	1.2	652	23	AAK75704	DNA encoding novel
809	18	1.2	654	11	AAQ04332	Plasmid pRCM1 and
810	18	1.2	657	22	AAK07935	Human TDC protein-
811	18	1.2	660	21	AAK96339	CDNA encoding a no
812	18	1.2	660	22	AAK61558	Human DNA encoding

813	18	1.2	664	23	AA572398	DNA encoding novel
814	18	1.2	665	22	AAH07112	Human cDNA clone (
815	18	1.2	666	22	AA684604	Human zrf4 protein
816	18	1.2	666	23	AA591699	DNA encoding novel
817	18	1.2	671	22	AA529154	CDNA encoding for
818	18	1.2	672	20	AA539595	Human HCMV repress
819	18	1.2	673	22	AAAL19259	Human breast cancer
820	18	1.2	675	22	AAK57397	Human immune/haema
821	18	1.2	676	19	AA662136	Leishmania donovan
822	18	1.2	676	18	AAV47556	Leishmania antigen
823	18	1.2	676	24	AA596020	Leishmania antigen
824	18	1.2	683	23	AAK64693	Human immune/haema
825	18	1.2	685	22	AA575708	DNA encoding novel
826	18	1.2	690	22	AAK87406	Human immune/haema
827	18	1.2	691	22	AAK74447	Human PRO16 nucleo
828	18	1.2	692	22	AAAL02626	Human reproductive
829	18	1.2	692	22	AAAD08205	Human secreted pro
830	18	1.2	693	20	AA233643	Human breast tumou
831	18	1.2	694	19	AAV34166	Human secreted pro
832	18	1.2	694	19	AAV34166	Human cDNA encodin
833	18	1.2	696	22	AA526254	Human breast cancer
834	18	1.2	699	22	ABAI7101	Human nervous syst
835	18	1.2	699	22	AAK59991	Human immune/haema
836	18	1.2	700	21	AA293153	Human CBCEP10 nuc
837	18	1.2	700	21	AA247519	CDNA sequence of w
838	18	1.2	700	22	AAH92268	Human inflammatory
839	18	1.2	700	22	AAH92269	Human inflammatory
840	18	1.2	700	22	AAH92617	Human inflammatory
841	18	1.2	700	22	AAH92618	Human inflammatory
842	18	1.2	700	22	AAH93072	Human inflammatory
843	18	1.2	706	21	AAAC74313	Human secreted pro
844	18	1.2	706	22	AAH64918	Human secreted pro
845	18	1.2	708	22	AAI20885	Human breast cancer
846	18	1.2	709	22	AAK85680	Human immune/haema
847	18	1.2	711	22	AAK85689	Human immune/haema
848	18	1.2	720	22	ABAO7921	Human ovarian and
849	18	1.2	720	22	AAAL03738	Human reproductive
850	18	1.2	720	22	AAK84566	Human immune/haema
851	18	1.2	720	22	AAH07620	Human cDNA clone (
852	18	1.2	721	20	AA588004	Human secreted pro
853	18	1.2	724	23	AA591434	DNA encoding novel
854	18	1.2	725	20	AAK37532	Human secreted pro
855	18	1.2	725	21	AAI95619	Human neuroblastom
856	18	1.2	726	21	AAAC76265	Human ORFX ORF1820
857	18	1.2	726	22	AA533105	DNA encoding human
858	18	1.2	727	20	AA520467	Human secreted pro
859	18	1.2	727	20	AA527247	CDNA encoding nove
860	18	1.2	729	21	AAFL3585	Aspergillus oryzae
861	18	1.2	736	23	AA586863	DNA encoding novel
862	18	1.2	737	22	AAAL1977	Human breast cancer
863	18	1.2	744	22	AAI95742	Human neuroblastom
864	18	1.2	749	22	AAH64931	Human secreted pro
865	18	1.2	750	20	AAZ16810	Human gene express
866	18	1.2	750	20	AAZ16487	Human gene express
867	18	1.2	751	22	AAK76824	Human immune/haema
868	18	1.2	753	20	AAZ00347	Nucleotide sequenc
869	18	1.2	753	22	AA589005	Human FATP3 coding
870	18	1.2	755	20	AAZ16214	Human gene express
871	18	1.2	757	20	AAZ16857	Human gene express
872	18	1.2	757	20	AAZ16858	Human gene express
873	18	1.2	757	22	AAH04118	Human cDNA clone (
874	18	1.2	758	22	AA526521	Human cDNA encodin
875	18	1.2	761	22	AAH70851	Human cervical can
876	18	1.2	763	22	AAAC6932	Nucleotide sequenc
877	18	1.2	769	22	AAAL13748	Human breast cancer
878	18	1.2	775	21	AAAC61101	Human cDNA encodin
879	18	1.2	775	22	AAI95644	Human neuroblastom
880	18	1.2	776	20	AAZ52955	Human prostate tum
881	18	1.2	777	21	AAAC99821	Human secreted pro
882	18	1.2	778	20	AAZ24830	Human secreted pro
883	18	1.2	781	22	AAAL24847	Human breast cancer
884	18	1.2	783	21	AAAC74316	Human secreted pro
885	18	1.2	783	22	AAK77935	Human immune/haema

Sequence encoding
Human branched pro
Corn branched chal
Human breast cancer
Human cDNA 5'-end
Human cDNA clone r
Canine hookworm ne
Human secreted pro
Human secreted pro
Flea serine protea
Arabidopsis thalia
Partial human SIC6
Human reproductive
Human reproductive
Human breast cancer
Human secreted pro
Human immune/haema
Human secreted pro
Arabidopsis thalia
Human cDNA encodin
Human secreted pro
Human immune/haema
Arabidopsis thalia
Human secreted pro
Human colon cancer
Human eosinocyte C
DNA encoding enviro
Sequence complemen
Human gene express
Human neuroblastom
Human secreted pro
Human secreted pro
Human secreted pro
Human immune/haema
Human immune/haema
Human neuroblastom
Soybean Wuschel (W
Human secreted pro
Human secreted pro
Human secreted pro
Human foetal cDNA,
Human secreted pro
Human neuroblastom
Human secreted pro
Human secreted pro
Human secreted pro
Proliferative glom
Staphylococcus aur
Aureobasidium pull
Maize glutathione-
Human reproductive
Human immune/haema
Arabidopsis thalia
Human secreted pro
Human immune/haema
Human DNA marker c
Human genomic DNA
Human secreted pro
Human secreted pro
Interferon gamma g
Alzheimer's disease
Human secreted pro
Human ORFX ORF2334
Human cardiovascular
Human cDNA encodin
C-terminal portion
Human secreted pro
Human secreted pro
Human secreted pro
DNA encoding novel

959	18	1.2	1061	21	AAA26423	Human secreted pro
960	18	1.2	1062	21	AAC69475	Human secreted pro
961	18	1.2	1066	22	AAD04001	Human protein tyro
962	18	1.2	1066	24	AAS62527	CDNA sequence #314
963	18	1.2	1070	24	AAS62649	CDNA sequence #436
964	18	1.2	1084	21	AAC59873	Human secreted pro
965	18	1.2	1089	21	AAC99677	Human secreted pro
966	18	1.2	1089	23	AAS87173	DNA encoding novel
967	18	1.2	1094	21	AAC59923	Human secreted pro
968	18	1.2	1098	21	AAC79709	Human secreted pro
969	18	1.2	1100	17	AAT40216	Sequence encoding
970	18	1.2	1101	21	AAC77676	Human cancer assoc
971	18	1.2	1102	19	AAV28734	Murine smooth musc
972	18	1.2	1102	22	AAD21105	Murine SM22 alpha
973	18	1.2	1102	22	AAD20564	Murine smooth musc
974	18	1.2	1102	22	AAB26685	Mouse smooth muscl
975	18	1.2	1105	22	ABA17630	Human nervous syst
976	18	1.2	1107	21	AAC77646	Human cancer assoc
977	18	1.2	1109	22	AAS40964	CDNA encoding nove
978	18	1.2	1111	21	AAC74346	Human secreted pro
979	18	1.2	1113	21	AAZ97105	Human secreted pro
980	18	1.2	1114	24	AAS62383	CDNA sequence #170
981	18	1.2	1121	22	AAD07933	Human TDC protein-
982	18	1.2	1125	22	ABA06543	Human CDNA SPQ ID
983	18	1.2	1130	22	AAZ24818	Human secreted pro
984	18	1.2	1130	22	AA106810	Human reproductive
985	18	1.2	1130	22	AA162691	Human breast or ov
986	18	1.2	1132	22	AA106808	Human reproductive
987	18	1.2	1132	22	AA162689	Human breast or ov
988	18	1.2	1137	22	AAK71677	Human immune/haema
989	18	1.2	1149	22	AAK32177	Human secreted pro
990	18	1.2	1160	21	AAC39976	Arabidopsis thalia
991	18	1.2	1161	24	AB101566	Human secreted pro
992	18	1.2	1165	18	AAT72945	Phaffia carotenoid
993	18	1.2	1167	22	AAF64199	Human secreted pro
994	18	1.2	1181	24	AAS62655	CDNA sequence #452
995	18	1.2	1189	20	AAZ00813	Human secreted pro
996	18	1.2	1201	21	AAC44975	Arabidopsis thalia
997	18	1.2	1205	21	AAC33121	Arabidopsis thalia
998	18	1.2	1210	24	ABA92619	Human class II ami
999	18	1.2	1228	21	AAC80567	Human secreted pro
1000	18	1.2	1238	22	AAK65390	Human immune/haema

ALIGNMENTS

RESULT 1

AH20992 standard; CDNA, 1539 BP.

XX	AAH20992;	
AC	31-AUG-2001 (first entry)	
XX		
DT		
XX	D. melanogaster acetylcholine receptor beta subunit Db3 cDNA.	
XX	Acetylcholine receptor beta subunit; Db3; fruitfly; insecticidal;	
KW	transgenic; modulator; plant-protection agent; medicine; ss.	
XX		
OS	Drosophila melanogaster.	
XX		
XX		
FT	key	Location/Qualifiers
FT	CDS	43..1368
FT		/tag= a
FT		/product= "Db3"
XX		
PN	EP1106689-A2.	
XX		
PD	13-JUN-2001.	
XX		
PF	28-NOV-2000; 2000EP-0125300.	
XX		

PR 10-DEC-1999; 99DE-1059582.
 XX (FARB) BAYER AG.
 PA Adamczewski M, Methfessel C, Schulte T;
 XX WPI; 2001-376906/40.
 DR P-PDB; AAB86218.
 XX New nucleic acid encoding Drosophila acetylcholine receptor subunit,
 PT useful in screening for plant protection agents and pharmaceuticals.
 PS Claim 1; Page 10-13; 18pp; German.
 XX
 CC This invention describes a novel nucleic acid (I) which encodes a
 CC Drosophila melanogaster acetylcholine receptor beta subunit, Db3 which
 CC has insecticidal activity. (II), related host cells, polypeptides, ACR,
 CC antibodies, transgenic invertebrates and regulatory regions, are used to
 CC identify agents that modulate ACR activity, potentially useful as
 CC plant-protection agents (e.g. insecticides) or pharmaceuticals for human
 CC or veterinary medicine. This sequence encodes the Drosophila melanogaster
 CC acetylcholine receptor beta subunit described in the method of the
 CC invention.
 XX
 SQ Sequence 1539 BP; 352 A; 476 C; 411 G; 300 T; 0 other;

Query Match 100.0%; Score 1539; DB 22; Length 1539;
 Best Local Similarity 100.0%; Pred. No. 0;
 Matches 1539; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY	1	atccgacagcaggtatcatccgaacaaagcgcgtgacaaatgacagactcccaag	60
DB	1	atccgacagcaggtatcatccgaacaaagcgcgtgacaaatgacagactcccaag	60
QY	61	ataaagcaccagttcccggtcctgacacactactgtcgaatgtcctaaaggagatg	120
DB	61	ataaagcaccagttcccggtcctgacacactactgtcgaatgtcctaaaggagatg	120
QY	121	ctcttatggggtcgtacatcctcgtgacagcgcacatgcccacgcggaacccaaagcc	180
DB	121	ctcttatggggtcgtacatcctcgtgacagcgcacatgcccacgcggaacccaaagcc	180
QY	181	aatgtcaaggccctgtgatcgcctccacgcgcgtgttcaagaactagacagcagatgtg	240
DB	181	aatgtcaaggccctgtgatcgcctccacgcgcgtgttcaagaactagacagcagatgtg	240
QY	241	caagcgggtgttccaaggaacccccaagcagtgctccttgaaatgtgtgacactacata	300
DB	241	caagcgggtgttccaaggaacccccaagcagtgctccttgaaatgtgtgacactacata	300
QY	301	gacatcgacaggttgaacgcaagcgtgacacccactgtcgtgataatctccgatggaga	360
DB	301	gacatcgacaggttgaacgcaagcgtgacacccactgtcgtgataatctccgatggaga	360
QY	361	gacgaggaagcgcgtgtgcaaacccgtcaacatagacaacatcagcagatcacttgaag	420
DB	361	gacgaggaagcgcgtgtgcaaacccgtcaacatagacaacatcagcagatcacttgaag	420
QY	421	tccagcgaagtcgtgaccccccaatcacacttccaagcgcagcagatgtgctcgaatg	480
DB	421	tccagcgaagtcgtgaccccccaatcacacttccaagcgcagcagatgtgctcgaatg	480
QY	481	gccgaacccaggttgacccctcagcagcagatgacacttccggtgagtgctcgaagcgtg	540
DB	481	gccgaacccaggttgacccctcagcagcagatgacacttccggtgagtgctcgaagcgtg	540
QY	541	tacacggtcactatgcgaactcaacatgctcaactggtccccaagcagacagagctgcaag	600
DB	541	tacacggtcactatgcgaactcaacatgctcaactggtccccaagcagacagagctgcaag	600
QY	601	ttgaagatcggtcctcgtggggtcgtgaaggtcgtcctcgtcggagagaagcagcaggaaga	660
DB	601	ttgaagatcggtcctcgtggggtcgtgaaggtcgtcctcgtcggagagaagcagcaggaaga	660

Db 601 ttgaagatcggtctcttggggccttgaaggtctcttgcggagaaagcagcgagagaga 660
 Qy gactcccttgaacacagcagccttggcttcagtcacccgagatgggaatcgttgaatcgcga 720
 Db 661 gattcccttgaacacagcagccttggcttcagtcacccgagatgggaatcgttgaatcgcga 720
 Qy 721 gcccaacttggatcagcagcactactacggtacatacgtgagatcacatctgaacggtcgaacg 780
 Db 721 gcccaacttggatcagcagcactactacggtacatacgtgagatcacatctgaacggtcgaacg 780
 Qy 781 cgtctcccatgtatacagcgccgtatctatacaccgcgcttccatcgtatctccttgcgc 840
 Db 781 cgtctcccatgtatacagcgccgtatctatacaccgcgcttccatcgtatctccttgcgc 840
 Qy 841 ctccacagccttctggcgcctcccccacatggcgcgagaaagatcatgataaagcgtc 900
 Db 841 ctccacagccttctggcgcctcccccacatggcgcgagaaagatcatgataaagcgtc 900
 Qy 901 ctcatcactgtatcagcgccttccatcgtatctcgtccacgctcctcctcagtgcttcc 960
 Db 901 ctcatcactgtatcagcgccttccatcgtatctcgtccacgctcctcctcagtgcttcc 960
 Qy 1081 gaagcgtctgagaaagcgtctgacagcgacacgtgacgctgctcgtctcgtctc 1140
 Db 1081 gaagcgtctgagaaagcgtctgacagcgacacgtgacgctgctcgtctcgtctcgtctc 1140
 Qy 1141 agcaccaactgtgagatcttctacagcaccagcctgtatctgagcgtctcc 1020
 Db 1141 agcaccaactgtgagatcttctacagcaccagcctgtatctgagcgtctcc 1020
 Qy 1201 gaagcgtgagaaagcgtctgacagcgacacgtgacgctgctcgtctcgtctcgtctc 1260
 Db 1201 gaagcgtgagaaagcgtctgacagcgacacgtgacgctgctcgtctcgtctcgtctcgtc 1260
 Qy 1261 aagcgtgagaaagcgtctgacagcgacacgtgacgctgctcgtctcgtctcgtctcgtc 1320
 Db 1261 aagcgtgagaaagcgtctgacagcgacacgtgacgctgctcgtctcgtctcgtctcgtc 1320
 Qy 1321 tctctcagcgtgacgtctcctcactatcttgcacatcaggtctcgtctgagagatcgtcga 1380
 Db 1321 tctctcagcgtgacgtctcctcactatcttgcacatcaggtctcgtctgagagatcgtcga 1380
 Qy 1381 ctcaagcgtgagaaagcgtctgacagcgacacgtgacgctgctcgtctcgtctcgtctcgtc 1440
 Db 1381 ctcaagcgtgagaaagcgtctgacagcgacacgtgacgctgctcgtctcgtctcgtctcgtc 1440
 Qy 1441 gttattatgtgtgtgacgacttaataatattatgagagacgtgagagagagagagagag 1500
 Db 1441 gttattatgtgtgtgacgacttaataatattatgagagacgtgagagagagagagagagag 1500
 Qy 1501 gacctctgcgaatgtctgtcttacaacacacacacacacacacacacacacacacacacac 1539
 Db 1501 gacctctgcgaatgtctgtcttacaacacacacacacacacacacacacacacacacacac 1539

RESULT 2

ABL23317 standard; DNA: 1362 BP.

ABL23317;

26-MAR-2002 (first entry)

Drosophila melanogaster genomic polynucleotide SEQ ID NO 21424.

KW Drosophila; developmental biology; cell signalling; insecticide;
 KM pharmaceutical; gene; ds.
 XX Drosophila melanogaster.
 OS WO200171042-A2.
 XX 27-SEP-2001.
 PD 23-MAR-2001; 2001WO-US09231.
 PF 23-MAR-2000; 2000US-191637P.
 PR 11-JUL-2000; 2000US-0614150.
 XX (PEKE) PE CORP NY.
 PA Venter JC, Adams M, Li PWD, Myers EW;
 PI WPI; 2001-656860/75.
 DR New isolated nucleic acid detection reagent for detecting 1000 or more
 PT genes from Drosophila and for elucidating cell signalling and cell-cell
 PT interactions -
 PS Claim 1: SEQ ID NO 21424; 21pp + Sequence Listing; English.
 XX
 CC The invention relates to an isolated nucleic acid detection reagent
 CC capable of detecting 1000 or more genes from Drosophila. The invention is
 CC useful in developmental biology and in elucidating cell signalling and
 CC cell-cell interactions in higher eukaryotes for the development of
 CC insecticides, therapeutics and pharmaceutical drugs. The invention
 CC discloses genomic DNA sequences (AB16176-AB130511), expressed DNA
 CC sequences (AB101840-AB16175) and the encoded proteins
 CC (AB161737-AB162072).
 CC The sequence data for this patent did not form part of the printed
 CC specification, but was obtained in electronic format directly from WIPO
 CC at ftp.wipo.int/pub/published_pct_sequences.
 CC
 XX
 SQ Sequence 1362 BP; 294 A; 444 C; 374 G; 250 T; 0 other;
 Query Match 78.1%; Score 1202; DB 23; Length 1362;
 Best Local Similarity 99.8%; Pred. No. 0;
 Matches 1352; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
 Qy 14 gtacatccgaacaaagcgcgtgaaacatgacgagactcccaagataaaggaccag 73
 Db 8 gtacatccgaacaaagcgcgtgaaacatgacgagactcccaagataaaggaccag 67
 Qy 74 tttccggtctgagactgacactactgtctgcaaatgcaatgagagatcttcttggggc 133
 Db 68 tttccggtctgagactgacactactgtctgcaaatgcaatgagagatcttcttggggc 127
 Qy 134 tgacttcgtgacagcgccacttgcacacgagaccccaagaacgcaatgtcaagggc 193
 Db 128 tgacttcgtgacagcgccacttgcacacgagaccccaagaacgcaatgtcaagggc 187
 Qy 194 tggatcgcttccacagcgccgtgttcaagaaactcagacagcagatgtcagcggtgtcc 253
 Db 188 tggatcgcttccacagcgccgtgttcaagaaactcagacagcagatgtcagcggtgtcc 247
 Qy 254 aaggaaccccccaagaacgtgtcccttggaatgtgtgtcgaactatagatgacacgaagt 313
 Db 248 aaggaaccccccaagaacgtgtcccttggaatgtgtgtcgaactatagatgacacgaagt 307
 Qy 314 tgaacggaacgttgaac 373
 Db 308 tgaacggaacgttgaac 367
 Qy 374 tgtgcaacggttcaatatagacaacatcagcagatcacttgaatcagcagaggtct 433
 Db 368 tgtgcaacggttcaatatagacaacatcagcagatcacttgaatcagcagaggtct 427

QY 434 ggaaccccaatcaactcttcaacgagcagaagtgctgatgagccgaacccag 493
 DB 428 ggaaccccaatcaactcttcaacgagcagaagtgctgatgagccgaacccag 487
 QY 494 tgaaccccaatcaactcttcaacgagcagaagtgctgatgagccgaacccag 553
 DB 488 tgaaccccaatcaactcttcaacgagcagaagtgctgatgagccgaacccag 547
 QY 554 ggaaccccaatcaactcttcaacgagcagaagtgctgatgagccgaacccag 613
 DB 548 ggaaccccaatcaactcttcaacgagcagaagtgctgatgagccgaacccag 607
 QY 614 cctggagcctgaagtgctctcgcggaagacgagcagagagagagagagagag 673
 DB 608 cctggagcctgaagtgctctcgcggaagacgagcagagagagagagagagag 667
 QY 674 agcagacacgctgctcagcagagagagagagagagagagagagagagagagag 733
 DB 668 agcagacacgctgctcagcagagagagagagagagagagagagagagagagag 727
 QY 734 gtaagactactacgctacatgagatcaactctgacgagcagcagcagcagcag 793
 DB 728 gtaagactactacgctacatgagatcaactctgacgagcagcagcagcagcag 787
 QY 794 acaagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcag 853
 DB 788 acaagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcag 847
 QY 854 ggtgctgctcagcagcagcagcagcagcagcagcagcagcagcagcagcagcag 913
 DB 848 ggtgctgctcagcagcagcagcagcagcagcagcagcagcagcagcagcagcag 907
 QY 914 tgcgagcctcctcagcagcagcagcagcagcagcagcagcagcagcagcagcag 973
 DB 908 tgcgagcctcctcagcagcagcagcagcagcagcagcagcagcagcagcagcag 967
 QY 974 ttgtgtaactcttcaacgagcagcagcagcagcagcagcagcagcagcagcag 1033
 DB 968 ttgtgtaactcttcaacgagcagcagcagcagcagcagcagcagcagcagcag 1027
 QY 1034 ttctagttctgtaactcttcaacgagcagcagcagcagcagcagcagcagcag 1093
 DB 1028 ttctagttctgtaactcttcaacgagcagcagcagcagcagcagcagcagcag 1087
 QY 1094 agctgtgtaactcttcaacgagcagcagcagcagcagcagcagcagcagcag 1153
 DB 1088 agctgtgtaactcttcaacgagcagcagcagcagcagcagcagcagcagcag 1147
 QY 1154 agtcgagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcag 1213
 DB 1148 agtcgagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcag 1207
 QY 1214 aggaagtcagcagcagcagcagcagcagcagcagcagcagcagcagcagcag 1273
 DB 1208 aggaagtcagcagcagcagcagcagcagcagcagcagcagcagcagcagcag 1267
 QY 1274 tgcagtgagcagcagcagcagcagcagcagcagcagcagcagcagcagcag 1333
 DB 1268 tgcagtgagcagcagcagcagcagcagcagcagcagcagcagcagcagcag 1327
 QY 1334 ccttctcctatcttgcagcagcagcagcagcagcagcagcagcagcagcag 1368
 DB 1328 ccttctcctatcttgcagcagcagcagcagcagcagcagcagcagcagcag 1362

RESULT 3

ABL23316/c

AC ABL23316;

DT 26-MAR-2002 (first entry)

XX

DE Drosophila melanogaster genomic polynucleotide SEQ ID NO 21421.
 XX Drosophila; developmental biology; cell signalling; insecticide;
 KW pharmaceutical; gene; ds.
 XX
 OS Drosophila melanogaster.
 PN WO200171042-A2.
 XX
 PD 27-SEP-2001.
 XX
 PF 23-MAR-2001; 2001WO-US09231.
 XX
 PR 23-MAR-2000; 2000US-191637P.
 PR 11-JUL-2000; 2000US-0614150.
 XX
 PA (PEKE) PE CORP NY.
 XX
 PI Venter JC, Adams M, Li PMD, Myers EW;
 XX
 DR MPI; 2001-656860/75.
 XX
 PT New isolated nucleic acid detection reagent for detecting 1000 or more
 PT genes from Drosophila and for elucidating cell signalling and cell-cell
 PT interactions -
 PS
 XX
 Claim 1: SEQ ID NO 21421; 21bp + Sequence Listing; English.
 CC The invention relates to an isolated nucleic acid detection reagent
 CC capable of detecting 1000 or more genes from Drosophila. The invention is
 CC useful in developmental biology and in elucidating cell signalling and
 CC cell-cell interactions in higher eukaryotes for the development of
 CC insecticides, therapeutics and pharmaceutical drugs. The invention
 CC discloses genomic DNA sequences (AB101840-AB16175), expressed DNA
 CC sequences (AB101840-AB16175) and the encoded proteins
 CC (AB16175-AB16175).
 CC The sequence data for this patent did not form part of the printed
 CC specification, but was obtained in electronic format directly from WIPO
 CC at ftp.wipo.int/pub/published_pat_sequences.
 CC
 SO Sequence 3798 BP; 869 A; 1013 C; 1048 G; 868 T; 0 other;
 Query Match 40.8%; Score 628; DB 23; Length 3798;
 Best Local Similarity 100.0%; Pred. No. 2.4e-253;
 Matches 628; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 352 cgaatgagagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 411
 DB 2138 CGATGGAG 2079
 QY 412 acttgaagtcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 471
 DB 2078 ACTTTGAAGTCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2019
 QY 472 ggcctgatgagcagcagcagcagcagcagcagcagcagcagcagcagcagc 531
 DB 2018 GGCCTGATGAGCGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1959
 QY 532 ccaagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcag 591
 DB 1958 CCAAGCGTGAACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1899
 QY 592 agctgcaagtcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 651
 DB 1898 AGCTGCAAGTTGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1839
 QY 652 gcaagagagagagcagcagcagcagcagcagcagcagcagcagcagcagcagcag 711
 DB 1838 GCGAG 1779
 QY 712 gactgagagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 771

PT New isolated nucleic acid detection reagent for detecting 1000 or more
PT genes from Drosophila and for elucidating cell signalling and cell-cell
PT interactions -
PS
PS Claim 1: SEQ ID NO 21409; 21bp + Sequence Listing; English.
XX
CC The invention relates to an isolated nucleic acid detection reagent
CC capable of detecting 1000 or more genes from Drosophila. The invention is
CC useful in developmental biology and in elucidating cell signalling and
CC cell-cell interactions in higher eukaryotes for the development of
CC insecticides, therapeutics and pharmaceutical drugs. The invention
CC discloses genomic DNA sequences (ABU16176-ABU30511), expressed DNA
CC sequences (ABU161840-ABU16175) and the encoded proteins
CC (ABU57737-ABU72072).
CC The sequence data for this patent did not form part of the printed
CC specification, but was obtained in electronic format directly from WIPO
CC at ftp.wipo.int/pub/published_pct_sequences.
XX
SQ Sequence 3764 BP; 943 A; 972 C; 960 G; 889 T; 0 other;

Query Match 12.7%; Score 195; DB 23; Length 3764;
Best Local Similarity 100.0%; Pred No. 4.5e-72;
Matches 195; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 157 gccaccgcgcgagccccaagaagcccaatgttcaagccctgctcaccgcgcgcctg 216
DB 271 GCCACGGCGGAGCCCAAGAACGCCAATGTCAAGCCCTGATGCCCTCACGGCGGCTG 212
QY 217 ttacgaacacagacagcagatgtgcagccggtgttccaagaagccccaagaacggttc 276
DB 211 TTACGAACTACGACAGCGATGTGACGCCGTTCACGAAGACCCCGACGATGTCC 152
QY 277 ctggaatggtgttcacccatcagacatgcagcaggttgaagcgaagctgaccaccac 336
DB 151 CTGGAATGTGTCTCATCATATGACATGAGAGATGTGAACGCCAAGCTGACCCAC 92
QY 337 tgctgctgaatc 351
DB 91 TGCTGCTGAATCTC 77

RESULT 6
AAL36376
ID AAL36376 standard; DNA; 571 BP.
XX
AC AAL36376;
XX
DE 08-JAN-2002 (first entry)
XX
DE Human musculoskeletal system related polynucleotide SEQ ID NO 2741.
XX
XX Cytostatic; immunosuppressive; nootropic; neuroprotective; antiviral;
XX anti-allergic; hepatotropic; antidiabetic; anti-inflammatory; antitumor;
XX anti-viral; anticonvulsant; antibacterial; antifungal; antiparasitic;
XX cardiac; gene therapy; cancer; immune disorder; cardiovascular disorder;
XX neurological disease; infection; human; secreted protein;
XX musculoskeletal system; ds.
XX
OS Homo sapiens.
XX
XX WO200155367-A1.
XX
PD 02-AUG-2001.
XX
PF 17-JAN-2001; 2001MO-US0138.
XX
XX 31-JAN-2000; 2000US-0179065.
XX 04-FEB-2000; 2000US-0180628.
XX 24-FEB-2000; 2000US-0184664.
XX 02-MAR-2000; 2000US-0186350.
XX 16-MAR-2000; 2000US-0189874.
XX 17-MAR-2000; 2000US-0190076.

PR 18-APR-2000; 2000US-0198123.
PR 19-MAY-2000; 2000US-0205515.
PR 07-JUN-2000; 2000US-0209467.
PR 28-JUN-2000; 2000US-0214886.
PR 30-JUL-2000; 2000US-0215135.
PR 07-JUL-2000; 2000US-0216647.
PR 07-JUL-2000; 2000US-0216880.
PR 11-JUL-2000; 2000US-0217487.
PR 11-JUL-2000; 2000US-0217496.
PR 14-JUL-2000; 2000US-0218290.
PR 26-JUL-2000; 2000US-0220963.
PR 26-JUL-2000; 2000US-0220964.
PR 14-AUG-2000; 2000US-0224518.
PR 14-AUG-2000; 2000US-0224519.
PR 14-AUG-2000; 2000US-0225213.
PR 14-AUG-2000; 2000US-0225214.
PR 14-AUG-2000; 2000US-0225266.
PR 14-AUG-2000; 2000US-0225267.
PR 14-AUG-2000; 2000US-0225268.
PR 14-AUG-2000; 2000US-0225270.
PR 14-AUG-2000; 2000US-0225447.
PR 14-AUG-2000; 2000US-0225757.
PR 14-AUG-2000; 2000US-0225758.
PR 14-AUG-2000; 2000US-0225759.
PR 18-AUG-2000; 2000US-0226279.
PR 22-AUG-2000; 2000US-0226681.
PR 22-AUG-2000; 2000US-0226688.
PR 22-AUG-2000; 2000US-0227182.
PR 23-AUG-2000; 2000US-0227009.
PR 30-AUG-2000; 2000US-0228924.
PR 01-SEP-2000; 2000US-0229287.
PR 01-SEP-2000; 2000US-0229343.
PR 01-SEP-2000; 2000US-0229344.
PR 01-SEP-2000; 2000US-0229345.
PR 05-SEP-2000; 2000US-0229509.
PR 05-SEP-2000; 2000US-0229513.
PR 06-SEP-2000; 2000US-0230437.
PR 06-SEP-2000; 2000US-0230438.
PR 08-SEP-2000; 2000US-0231242.
PR 08-SEP-2000; 2000US-0231243.
PR 08-SEP-2000; 2000US-0231244.
PR 08-SEP-2000; 2000US-0231413.
PR 08-SEP-2000; 2000US-0231414.
PR 08-SEP-2000; 2000US-0232080.
PR 08-SEP-2000; 2000US-0232081.
PR 12-SEP-2000; 2000US-0231968.
PR 14-SEP-2000; 2000US-0232397.
PR 14-SEP-2000; 2000US-0232398.
PR 14-SEP-2000; 2000US-0232399.
PR 14-SEP-2000; 2000US-0232400.
PR 14-SEP-2000; 2000US-0232401.
PR 14-SEP-2000; 2000US-0233063.
PR 14-SEP-2000; 2000US-0233064.
PR 14-SEP-2000; 2000US-0233065.
PR 21-SEP-2000; 2000US-0234223.
PR 21-SEP-2000; 2000US-0234274.
PR 25-SEP-2000; 2000US-0234997.
PR 25-SEP-2000; 2000US-0234998.
PR 26-SEP-2000; 2000US-0235484.
PR 27-SEP-2000; 2000US-0235834.
PR 27-SEP-2000; 2000US-0235836.
PR 29-SEP-2000; 2000US-0236327.
PR 29-SEP-2000; 2000US-0236367.
PR 29-SEP-2000; 2000US-0236368.
PR 29-SEP-2000; 2000US-0236369.
PR 29-SEP-2000; 2000US-0236370.
PR 02-OCT-2000; 2000US-0236802.
PR 02-OCT-2000; 2000US-0237037.
PR 02-OCT-2000; 2000US-0237038.
PR 02-OCT-2000; 2000US-0237039.
PR 02-OCT-2000; 2000US-0237040.
PR 13-OCT-2000; 2000US-0239935.
PR 13-OCT-2000; 2000US-0239937.

PR 20-OCT-2000; 2000US-0240960.
PR 20-OCT-2000; 2000US-0241221.
PR 20-OCT-2000; 2000US-0241785.
PR 20-OCT-2000; 2000US-0241785.
PR 20-OCT-2000; 2000US-0241786.
PR 20-OCT-2000; 2000US-0241787.
PR 20-OCT-2000; 2000US-0241808.
PR 20-OCT-2000; 2000US-0241809.
PR 20-OCT-2000; 2000US-0241826.
PR 01-NOV-2000; 2000US-0244617.
PR 08-NOV-2000; 2000US-0246474.
PR 08-NOV-2000; 2000US-0246475.
PR 08-NOV-2000; 2000US-0246476.
PR 08-NOV-2000; 2000US-0246477.
PR 08-NOV-2000; 2000US-0246478.
PR 08-NOV-2000; 2000US-0246523.
PR 08-NOV-2000; 2000US-0246524.
PR 08-NOV-2000; 2000US-0246525.
PR 08-NOV-2000; 2000US-0246526.
PR 08-NOV-2000; 2000US-0246527.
PR 08-NOV-2000; 2000US-0246528.
PR 08-NOV-2000; 2000US-0246532.
PR 08-NOV-2000; 2000US-0246609.
PR 08-NOV-2000; 2000US-0246610.
PR 08-NOV-2000; 2000US-0246611.
PR 08-NOV-2000; 2000US-0246613.
PR 17-NOV-2000; 2000US-0249207.
PR 17-NOV-2000; 2000US-0249208.
PR 17-NOV-2000; 2000US-0249209.
PR 17-NOV-2000; 2000US-0249210.
PR 17-NOV-2000; 2000US-0249211.
PR 17-NOV-2000; 2000US-0249212.
PR 17-NOV-2000; 2000US-0249213.
PR 17-NOV-2000; 2000US-0249214.
PR 17-NOV-2000; 2000US-0249215.
PR 17-NOV-2000; 2000US-0249216.
PR 17-NOV-2000; 2000US-0249217.
PR 17-NOV-2000; 2000US-0249218.
PR 17-NOV-2000; 2000US-0249244.
PR 17-NOV-2000; 2000US-0249245.
PR 17-NOV-2000; 2000US-0249264.
PR 17-NOV-2000; 2000US-0249265.
PR 17-NOV-2000; 2000US-0249297.
PR 17-NOV-2000; 2000US-0249299.
PR 17-NOV-2000; 2000US-0249300.
PR 01-DEC-2000; 2000US-0250160.
PR 01-DEC-2000; 2000US-0250391.
PR 05-DEC-2000; 2000US-0251030.
PR 05-DEC-2000; 2000US-0251988.
PR 05-DEC-2000; 2000US-0256719.
PR 06-DEC-2000; 2000US-0251479.
PR 08-DEC-2000; 2000US-0251856.
PR 08-DEC-2000; 2000US-0251868.
PR 08-DEC-2000; 2000US-0251869.
PR 08-DEC-2000; 2000US-0251989.
PR 08-DEC-2000; 2000US-0251990.
PR 11-DEC-2000; 2000US-0254097.
PR 05-JAN-2001; 2001US-0259678.
XX
XX (HUMA-) HUMAN GENOME SCI INC.
XX
XX Rosen CA, Barash SC, Ruben SM;
XX
XX WPI; 2001-451937/48.
XX
XX
XX Isolated polypeptide for treating, preventing and/ or prognosing
XX disorders related to the musculoskeletal system including
XX musculoskeletal cancers and also for testing and detection e.g.
XX diagnosis -
XX
XX Example 2; SEQ ID NO 2741; 781pp + Sequence Listing; English.
XX
XX The invention relates to novel genes (PAL34669-PAL37666) and proteins
XX (ABB03087-ABB04109) associated with the musculoskeletal system useful
CC

CC for preventing, treating or ameliorating medical conditions e.g. by
CC protein or gene therapy. The genes are isolated from a range of human
CC tissues disclosed in the specification. The nucleic acids, proteins,
CC antibodies and (ant)agonists are useful in the diagnosis, treatment
CC and prevention of: (a) cancer, e.g. breast and ovarian cancer and
CC other cancers of the adrenal gland, bone, bone marrow, breast,
CC gastrointestinal tract, liver, lung, or urogenital; (b) immune
CC disorders e.g. Addison's disease, allergies, autoimmune haemolytic
CC anaemia, autoimmune thyroiditis, diabetes mellitus, Crohn's disease,
CC multiple sclerosis, rheumatoid arthritis and ulcerative colitis;
CC (c) cardiovascular disorders such as atherosclerotic ischaemias; (d) wound
CC healing; (e) neurological diseases e.g. cerebral anoxia and epilepsy;
CC and (f) infectious diseases such as viral, bacterial, fungal and
CC parasitic infections.
CC Note: The sequence data for this patent did not form part of the
CC printed specification, but was obtained in electronic format directly
CC from WIPO at ftp.wipo.int/pub/published_pat_sequences.
XX

Query Match	1.6%	Score 24	DB 22	Length 571
Best Local Similarity	100.0%	Pred. No. 2		
Matches	24	Conservative	0	Mismatches 0; Indels 0; Gaps 0;
Oy	1516	tctgcttacaataaaaaaaaaa	1539	
db	490	tctgcttacaataaaaaaaaaa	513	

RESULT	7
AAH41137	
ID	AAH41137 standard; DNA; 1135 BP.

DT 21-AUG-2001 (first entry)

DE Arabidopsis thaliana chlorophyllase coding sequence #2

OS Arabidopsis thaliana.

PN JP2001086990-A.

PD 03-APR-2001.

PF 20-SEP-1999; 99JP-0266181.

PR 20-SEP-1999; 99JP-0266181.

PA (KAGO) KAGOME KK.

DR WPI; 2001-338421/36.

XX

XX

XX

CC chlorophyllase protein coding sequence can be used for the

[illegible]

Query Match	Score	DB	Length
1.48;	22;	22;	1135;

Matches	22;	Conservative	0;	Mismatches	0;	Indels	0;	Gaps	0;
---------	-----	--------------	----	------------	----	--------	----	------	----

QY	1518	tgcttcaaaaaaaaaaaaaa	1539
----	------	----------------------	------

Db 1106 Tgcttaccatacaaaaaaaaaa 1127

RESULT 8

AAZ24886
ID AAZ24886 standard; DNA; 1377 BP.

XX AAZ24886;

DT 02-DEC-1999 (first entry)

DE Human secreted protein gene 76 clone HFPBM30.

XX Human; secreted protein; fusion protein; gene therapy; protein therapy;
KW diagnosis; tissue; cancer; tumour; neurodegenerative disorder; leukaemia;
KW developmental abnormality; foetal deficiency; blood; allergy; renal; ds;
KW immune system; asthma; lymphocytic disease; brain; hepatic; lymphoma;
KW inflammation; ischaemic shock; Alzheimer's disease; restenosis; AIDS;
KW cognitive disorder; schizophrenia; prostate; obesity; osteoclast; thymus;
KW osteoporosis; arthritis; testis; lung; thyroiditis; thyroid; digestion;
KW endocrine; metabolism; regulation; malabsorption; gastritis; neoplasm.

OS Homo sapiens.

PN WO9947540-A1.

PD 23-SEP-1999.

PF 18-MAR-1999; 99WO-US05804.

XX 19-MAR-1998; 98US-0078563.

PR 19-MAR-1998; 98US-0078566.

PR 19-MAR-1998; 98US-0078573.

PR 19-MAR-1998; 98US-0078574.

PR 19-MAR-1998; 98US-0078576.

PR 19-MAR-1998; 98US-0078577.

PR 19-MAR-1998; 98US-0078578.

PR 19-MAR-1998; 98US-0078579.

PR 01-APR-1998; 98US-0080312.

PR 01-APR-1998; 98US-0080313.

PR 01-APR-1998; 98US-0080314.

XX (HUMA-) HUMAN GENOME SCI INC.

PI Ruben SM, Ni J, Rosen CA, Yu G, Young PE, Feng P, Soppet DR;

PI Wei Y, Endress GA, Duan RD, Kyaw H, Ebner R, Lafleur DM;

PI Olsen HS, Shi Y, Moore PA;

DR WPI; 1999-562050/47.

DR P-PSDB; AAY41383.

XX New isolated human genes, useful for diagnosis and treatment of e.g.

PT cancers, neurological disorders, immune diseases, inflammation or blood

PT disorders

PS Claim 1; Page 343-344; 484pp; English.

XX This sequence represents a nucleic acid molecule which encodes a

CC secreted human protein. The gene number and the clone it is derived

CC from, are detailed in the descriptor line. The gene can be used to

CC generate fusion proteins by linking to the gene to a human immunoglobulin

CC Fc portion (e.g. AAZ24802) for increasing the stability of the fused

CC protein as compared to the human protein only.

CC The invention relates to 95 novel genes and their fragments (nucleic

CC acid sequences: AAZ24811-224907; amino acid sequences AAY41308-741404)

CC which are useful for preventing, treating or ameliorating medical

CC conditions e.g. by protein or gene therapy. Also, pathological

CC conditions can be diagnosed by determining the amount of the new

CC polypeptides in a sample or by determining the presence of mutations in

CC the new polynucleotides. Specific uses are described for each of the 95

CC polynucleotides, based on which tissues they are most highly expressed in

CC (see AAZ24811 for described uses).

XX Sequence 1377 BP; 311 A; 283 C; 379 G; 404 T; 0 other;

SO Query Match 1.4%; Score 21; DB 20; Length 1377;

Best Local Similarity 100.0%; Pred. No. 33;

Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1519 gcttaccatacaaaaaaaaaa 1539

DB 1354 gcttaccatacaaaaaaaaaa 1374

RESULT 9

AAH32948
ID AAH32948 standard; cDNA; 1843 BP.

XX AAH32948;

DT 03-SEP-2001 (first entry)

DE Human colon cancer antigen encoding cDNA SEQ ID NO:4.

XX Human; colon cancer; colon cancer antigen; diagnosis; detection;

KW colorectal carcinoma; ss.

OS Homo sapiens.

PN WO200122920-A2.

XX 05-APR-2001.

PF 28-SEP-2000; 2000WO-US26524.

PR 29-SEP-1999; 99US-0157137.

PR 03-NOV-1999; 99US-0163280.

XX (HUMA-) HUMAN GENOME SCI INC.

PI Ruben SM, Barash SC, Birse CE, Rosen CA;

PI WPI; 2001-235357/24.

PI P-PSDB; AAG73517.

XX Nucleic acids encoding 4277 human colon cancer-associated polypeptides,

XX useful for preventing, diagnosing and/or treating colorectal cancers -

XX Claim 1; Page 2231-2232; 9803pp; English.

XX AAH32943 to AAH37195 and AAG73514 to AAG77788 represent human colon

XX cancer-associated nucleic acid molecules (N) and proteins (P), where

XX the proteins are collectively known as colon cancer antigens. The colon

XX cancer antigens have cytostatic activity and can be used in gene

XX therapy and vaccine production. N and P may be used in the prevention,

XX diagnosis and treatment of diseases associated with inappropriate P

XX expression. For example, N and P may be used to treat disorders

XX associated with decreased expression by rectifying mutations or deletions

XX in a patient's genome that affect the activity of P by expressing

XX inactive proteins or to supplement the patient's own production of P.

XX Additionally, N may be used to produce the colon cancer-associated P,

XX by inserting the nucleic acids into a host cell and culturing the cell

XX to express the proteins. N and P can be used in the prevention, diagnosis

XX and treatment of colorectal carcinomas and cancers. AAH37196 to AAH37204

XX and AAG77789 represent sequences used in the exemplification of the

XX present invention.

XX N.B. Pages 666 to 682 and page 7053 of the sequence listing were

XX missing at time of publication, meaning no sequences are present for

XX SEQ ID NO:1027 to 1052, 7921 and 7922.

XX Sequence 1843 BP; 632 A; 329 C; 310 G; 562 T; 10 other;

Query Match 1.4%; Score 21; DB 22; Length 1843;
Best Local Similarity 100.0%; Pred. No. 33;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1519 gcttacaataaaaaaa 1538
|||||
DB 1746 gcttacaataaaaaaa 1766

RESULT 10

ABL24926
ID ABL24926 standard; DNA; 4331 BP.

XX ABL24926;
XX

DT 26-MAR-2002 (first entry)

DE Drosophila melanogaster genomic polynucleotide SEQ ID NO 26251.

XX Drosophila; developmental biology; cell signalling; insecticide;
XX pharmaceutical; gene; ds.

OS Drosophila melanogaster.

XX WO200171042-A2.

XX 27-SEP-2001.

XX 23-MAR-2001; 2001WO-US09231.

XX 23-MAR-2000; 2000US-191637P.

PR 11-JUL-2000; 2000US-0614150.

XX (PEKE) PE CORP NY.

XX Venter JC, Adams M, Li PWD, Myers EW;

DR WPI; 2001-656860/75.

PT New isolated nucleic acid detection reagent for detecting 1000 or more
PT genes from Drosophila and for elucidating cell signalling and cell-cell
PT interactions -

XX Claim 1; SEQ ID NO 26251; 21pp + Sequence Listing; English.

CC The invention relates to an isolated nucleic acid detection reagent
CC capable of detecting 1000 or more genes from Drosophila. The invention is
CC useful in developmental biology and in elucidating cell signalling and
CC cell-cell interactions in higher eukaryotes for the development of
CC insecticides, therapeutics and pharmaceutical drugs. The invention
CC discloses genomic DNA sequences (AB116176-AB130511), expressed DNA
CC sequences (AB101840-AB116175) and the encoded proteins
CC (ABB57737-ABB72072).

CC The sequence data for this patent did not form part of the printed
CC specification, but was obtained in electronic format directly from WIPO
CC at ftp.wipo.int/pub/published_pct_sequences.

XX Sequence 4331 BP; 1202 A; 1172 C; 1044 G; 913 T; 0 other;

Query Match 1.4%; Score 21; DB 23; Length 4331;
Best Local Similarity 100.0%; Pred. No. 30;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1518 tgcttacaataaaaaaa 1538
|||||
DB 3697 tgcttacaataaaaaaa 3717

RESULT 11

ABL31424
ID AAL31424 standard; DNA; 51 BP.

AC AAL31424;
XX
XX 24-JAN-2002 (first entry)

DE Human SNP oligonucleotide #4632.

XX Immunosuppressive; immunostimulatory; antiinflammatory; cytostatic;
XX neuroprotective; antimicrobial; gene therapy; vaccine; amylose; cancer;
XX amyloid protein; angiotensin; apoptosis related protein; cadherin;
XX cyclin; polymerase; oncogene; histone; kinase; colony stimulating factor;
XX complement related protein; cytochrome; kinesin; cytokine; interferon;
XX interleukin; G-protein coupled receptor; thioesterase; inflammation;
XX multifactorial disease; autoimmune disease; infection;
XX nervous system disease; ss.

XX Homo sapiens.

XX WO200147944-A2.

XX 05-JUL-2001.

XX 28-DEC-2000; 2000WO-US35498.

XX 28-DEC-1999; 99US-0173419.

PR 27-DEC-2000; 2000US-0173419.

XX (CURA-) CURAGEN CORP.

XX Shimkets RA, Leach M;

DR WPI; 2001-465210/50.

PT Polymorphic nucleic acids encoding e.g. amylases, cyclins, polymerases,
PT oncogenes and histones, useful for diagnosing and treating, e.g.
PT cancer, autoimmune diseases and infections -

XX Claim 1; Page 2718; 4143pp; English.

CC The present invention relates to oligonucleotides encoding polymorphic
CC variants of proteins related to amylases, amyloid proteins, angiotensin,
CC apoptosis related proteins, cadherin, cyclin, polymerase, oncogenes,
CC histones, kinases, colony stimulating factors, complement related
CC G-protein coupled receptors and thioesterases. Interferons, interleukins,
CC one such oligonucleotide. The oligonucleotides and the peptides encoded
CC by them may be used in the prevention, diagnosis and treatment of
CC diseases associated with inappropriate expression of the proteins listed
CC above. Disorders that may be prevented, diagnosed and/or treated include
CC multifactorial diseases with a genetic component, such as autoimmune
CC diseases (e.g. rheumatoid arthritis, multiple sclerosis, diabetes,
CC systemic lupus erythematosus and Grave's disease), inflammation, cancer
CC (e.g. cancers of the bladder, brain, breast, colon and kidney,
CC leukaemia), diseases of the nervous system and an infection of pathogenic
CC organisms.

XX Sequence 51 BP; 21 A; 9 C; 7 G; 14 T; 0 other;

Query Match 1.3%; Score 20; DB 22; Length 51;
Best Local Similarity 100.0%; Pred. No. 11e+02;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1519 gcttacaataaaaaaa 1538
|||||
DB 7 gcttacaataaaaaaa 26

RESULT 12

ABL31427
ID AAL31427 standard; DNA; 51 BP.

DT 24-JAN-2002 (first entry)
XX
DE Human SNP oligonucleotide #4635.
XX
KW Immunosuppressive; immunostimulatory; antiinflammatory; cytostatic;
KW neuroprotective; antimicrobial; gene therapy; vaccine; amylase; cancer;
KW amyloid protein; angiotensin; apoptosis related protein; cadherin;
KW cyclin; polymerase; oncogene; histone; kinase; colony stimulating factor;
KW complement related protein; cytochrome; kinesin; cytokine; interferon;
KW interleukin; G-protein coupled receptor; thioesterase; inflammation;
KW multifactorial disease; autoimmune disease; infection;
KW nervous system disease; ss.
XX
OS Homo sapiens.
XX
PN WO200147944-A2.
XX
PD 05-JUL-2001.
XX
PF 28-DEC-2000; 2000WO-US35498.
XX
PR 28-DEC-1999; 99US-0173419.
XX
PR 27-DEC-2000; 2000US-0173419.
XX
PA (CURA-) CURAGEN CORP.
XX
PI Shinkets RA, Leach M;
XX
DR WPI; 2001-465210/50.
XX
PT Polymorphic nucleic acids encoding e.g. amylases, cyclins, polymerases,
XX
PT oncogenes and histones, useful for diagnosing and treating, e.g.
XX
PT cancer, autoimmune diseases and infections -
XX
PS Claim 1; Page 2719; 4143pp; English.
XX
CC The present invention relates to oligonucleotides encoding polymorphic
CC variants of proteins related to amylases, amyloid proteins, angiotensin,
CC apoptosis related proteins, cadherin, cyclin, polymerase, oncogenes,
CC histones, kinases, colony stimulating factors, complement related
CC proteins, cytochromes, kinesins, cytokines, interferons, interleukins,
CC G-protein coupled receptors and thioesterases. The present sequence is
CC one such oligonucleotide. The oligonucleotides and the peptides encoded
CC by them may be used in the prevention, diagnosis and treatment of
CC diseases associated with inappropriate expression of the proteins listed
CC above. Disorders that may be prevented, diagnosed and/or treated include
CC multifactorial diseases with a genetic component, such as autoimmune
CC diseases (e.g. rheumatoid arthritis, multiple sclerosis, diabetes,
CC systemic lupus erythematosus and Grave's disease), inflammation, cancer
CC (e.g. cancers of the bladder, brain, breast, colon and kidney,
CC leukaemia), diseases of the nervous system and an infection of pathogenic
CC organisms.
XX
CC
SQ Sequence 51 BP; 26 A; 7 C; 8 G; 10 T; 0 other;
XX
XX
Query Match 1.3%; Score 20; DB 22; Length 51;
Best Local Similarity 100.0%; Pred. No. 1.1e+02;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1519 gcttacaataaaaaaaa 1538
DB 20 gcttacaataaaaaaaa 39

DE Human SNP oligonucleotide #5074.
XX
KW Immunosuppressive; immunostimulatory; antiinflammatory; cytostatic;
KW neuroprotective; antimicrobial; gene therapy; vaccine; amylase; cancer;
KW amyloid protein; angiotensin; apoptosis related protein; cadherin;
KW cyclin; polymerase; oncogene; histone; kinase; colony stimulating factor;
KW complement related protein; cytochrome; kinesin; cytokine; interferon;
KW interleukin; G-protein coupled receptor; thioesterase; inflammation;
KW multifactorial disease; autoimmune disease; infection;
KW nervous system disease; ss.
XX
OS Homo sapiens.
XX
PN WO200147944-A2.
XX
PD 05-JUL-2001.
XX
PF 28-DEC-2000; 2000WO-US35498.
XX
PR 28-DEC-1999; 99US-0173419.
XX
PR 27-DEC-2000; 2000US-0173419.
XX
PA (CURA-) CURAGEN CORP.
XX
PI Shinkets RA, Leach M;
XX
DR WPI; 2001-465210/50.
XX
PT Polymorphic nucleic acids encoding e.g. amylases, cyclins, polymerases,
XX
PT oncogenes and histones, useful for diagnosing and treating, e.g.
XX
PT cancer, autoimmune diseases and infections -
XX
PS Claim 1; Page 2846; 4143pp; English.
XX
CC The present invention relates to oligonucleotides encoding polymorphic
CC variants of proteins related to amylases, amyloid proteins, angiotensin,
CC apoptosis related proteins, cadherin, cyclin, polymerase, oncogenes,
CC histones, kinases, colony stimulating factors, complement related
CC proteins, cytochromes, kinesins, cytokines, interferons, interleukins,
CC G-protein coupled receptors and thioesterases. The present sequence is
CC one such oligonucleotide. The oligonucleotides and the peptides encoded
CC by them may be used in the prevention, diagnosis and treatment of
CC diseases associated with inappropriate expression of the proteins listed
CC above. Disorders that may be prevented, diagnosed and/or treated include
CC multifactorial diseases with a genetic component, such as autoimmune
CC diseases (e.g. rheumatoid arthritis, multiple sclerosis, diabetes,
CC systemic lupus erythematosus and Grave's disease), inflammation, cancer
CC (e.g. cancers of the bladder, brain, breast, colon and kidney,
CC leukaemia), diseases of the nervous system and an infection of pathogenic
CC organisms.
XX
CC
SQ Sequence 51 BP; 9 A; 5 C; 5 G; 32 T; 0 other;
XX
XX
Query Match 1.3%; Score 20; DB 22; Length 51;
Best Local Similarity 100.0%; Pred. No. 1.1e+02;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1520 cttaacaaaaaa 1539
DB 30 cttaacaaaaaa 11

RESULT 14
AAS56571
ID AAS56571 standard; cDNA; 217 BP.
XX
AC AAS56571;
XX
DT 18-DEC-2001 (first entry)
XX
DE Human cDNA for an ovarian cancer protein #195.
XX

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: July 18, 2002, 10:18:12 ; Search time 132.17 Seconds
(without alignments)
577.217 Million cell updates/sec

Title: US-09-732-680A-2
Perfect score: 2311
Sequence: 1 MTTTPKIKAPVSGPLDLL.....RISFVSFLAFLILAIRCSV 441

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 562222 seqs, 172994929 residues
Total number of hits satisfying chosen parameters: 562222

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

SPTRMBL_19:*
1: sp_archaea:*
2: sp_bacteria:*
3: sp_fungi:*
4: sp_human:*
5: sp_invertebrate:*
6: sp_mammal:*
7: sp_mhc:*
8: sp_organelle:*
9: sp_plant:*
10: sp_plant:*
11: sp_rodent:*
12: sp_virus:*
13: sp_vertebrate:*
14: sp_unclassified:*
15: sp_virus:*
16: sp_bacteriap:*
17: sp_archaeap:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match Length	ID	Description
1	2311	100.0	441 5 Q9VP08	Q9VP08 drosophila
2	1650	71.4	327 5 Q9SR44	Q9SR44 drosophila
3	434	18.8	512 11 Q91X60	Q91X60 mus musculu
4	418.5	18.1	457 6 Q9XS62	Q9XS62 canis famli
5	417	18.0	452 13 Q91RC7	Q91RC7 gallus gall
6	417	18.0	629 11 Q9ER51	Q9ER51 mus musculu
7	412	17.8	629 11 Q923N8	Q923N8 mus musculu
8	412	17.8	641 11 Q923N7	Q923N7 mus musculu
9	411.5	17.8	664 5 Q9NS08	Q9NS08 caenorhabdi
10	406	17.6	484 4 Q9BS29	Q9BS29 homo sapien
11	404.5	17.5	447 11 Q9JLB5	Q9JLB5 rattus norv
12	398	17.2	494 11 Q9R0M9	Q9R0M9 mus musculu
13	397.5	17.2	502 5 Q96AN6	Q96AN6 caenorhabdi
14	395	17.1	523 5 Q46128	Q46128 heliothis v
15	392.5	17.0	501 4 Q96F88	Q96F88 homo sapien
16	389	16.8	461 5 P91197	P91197 caenorhabdi

17	388.5	16.8	537 5 Q9U941	Q9U941 myzus persi
18	386	16.7	795 5 Q18394	Q18394 drosophila
19	385.5	16.7	545 5 Q96631	Q96631 heliothis v
20	382.5	16.6	515 5 Q46133	Q46133 locusta mig
21	382	16.5	552 5 P91765	P91765 myzus persi
22	380	16.4	484 11 Q70213	Q70213 cavia porce
23	379	16.4	568 5 Q9NFR5	Q9NFR5 drosophila
24	378.5	16.4	510 4 Q60854	Q60854 homo sapien
25	378	16.4	502 11 Q9JHD6	Q9JHD6 mus musculu
26	377	16.3	490 11 Q70212	Q70212 cavia porce
27	377	16.3	505 4 Q9BQ93	Q9BQ93 homo sapien
28	374.5	16.2	542 5 Q18556	Q18556 caenorhabdi
29	374.5	16.2	776 5 Q44202	Q44202 drosophila
30	372	16.1	502 4 Q96RH2	Q96RH2 homo sapien
31	371.5	16.1	489 4 Q9BRH4	Q9BRH4 homo sapien
32	371.5	16.1	531 5 Q96632	Q96632 heliothis v
33	370.5	16.0	489 4 Q96RH3	Q96RH3 homo sapien
34	370	16.0	483 6 Q9N0F4	Q9N0F4 mustela put
35	368	15.9	515 5 Q16926	Q16926 caenorhabdi
36	367	15.9	484 13 Q9PMS8	Q9PMS8 gallus gall
37	365.5	15.8	595 5 P91764	P91764 myzus persi
38	365	15.8	580 5 Q9VC72	Q9VC72 drosophila
39	364	15.8	477 11 Q62999	Q62999 rattus norv
40	362	15.7	489 11 Q61225	Q61225 mus musculu
41	359.5	15.6	432 11 Q9QXK6	Q9QXK6 mus musculu
42	358.5	15.5	772 5 Q9W3G6	Q9W3G6 drosophila
43	357.5	15.5	560 5 Q62083	Q62083 caenorhabdi
44	356	15.4	532 5 Q9U940	Q9U940 myzus persi
45	356	15.4	573 5 Q9GQ09	Q9GQ09 caenorhabdi

ALIGNMENTS

RESULT	ID	Q9VP08	PRELIMINARY:	PRT:	441 AA.
AC	Q9VP08	Q9VP08			
DT	01-MAY-2000	(TREMBLrel. 13. Created)			
DT	01-MAY-2000	(TREMBLrel. 13. Last sequence update)			
DT	01-DEC-2001	(TREMBLrel. 19. Last annotation update)			
DE	CG11822 PROTEIN (NICOTINIC ACETYLCHOLINE RECEPTOR BETA 3 (DBETA3))				
DE	SUBUNIT (NICOTINIC ACETYLCHOLINE RECEPTOR BETA SUBUNIT).				
GN	CG11822 OR NACRBETA-21C.				
OS	Drosophila melanogaster (Fruit fly).				
OC	Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;				
OC	Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;				
OC	Ephydroidea; Drosophilidae; Drosophila.				
OX	NCBI_TaxID=7227;				
RN	[1]				
RP	SEQUENCE FROM N.A.				
KC	STRAIN=BERKELEY;				
KX	MEDLINE=20196006; PubMed=10731132;				
RA	Adams M.D., Celinker S.E., Holt R.A., Evans C.A., Gocayne J.D.,				
RA	Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galie R.F.,				
RA	Georgopoulos R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,				
RA	Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,				
RA	Branden R.C., Rogers T.-H.C., Blazey R.G., Champe M., Pfeiffer B.D.,				
RA	Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,				
RA	Abriil J.F., Agbayani A., An H.-J., Andrews-Pfannkoch C., Baldwin D.,				
RA	Ballev R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,				
RA	Besson K.Y., Berooz P.V., Bernier B.P., Bhandari D., Bolshakov S.,				
RA	Botkova D., Botchan M.R., Bouck J., Brockstein P., Brothier P.,				
RA	Buttis K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,				
RA	Cherry J.M., Crawley S., Dahlke C., Davenport L.B., Davies P.,				
RA	de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,				
RA	Dodson K.K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,				
RA	Dubin K.J., Evangelista C.C., Ferraz C., Ferrera S., Fleischmann W.,				
RA	Foster C., Gabriellian A.E., Garg N.S., Gelbart W.M., Glasser K.,				
RA	Glodek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,				
RA	Harris N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck J.,				
RA	Hoslin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwam C.,				
RA	Jaitani M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,				

RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
 RA Lasko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,
 RA Liu X., Mattei B., McIntosh T.C., Moleod M.P., Moperson D.,
 RA Merkulov G., Milshina N.V., Mobarry C., Morris J., Moshfegh A.,
 RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
 RA Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Pacled J.M.,
 RA Palazzolo M., Pflum G.S., Pan S., Pollard J., Puri V., Reese M.G.,
 RA Reinert K., Remington K., Saunders R.D.C., Schaefer F., Shen H.,
 RA Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,
 RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,
 RA Strickas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
 RA Wang Z.-Y., Wasserman D.A., Weinstein G.M., Weissenbach J.,
 RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,
 RA Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
 RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,
 RA Glibis R.A., Myers E.W., Rubin G.M., Venter J.C.;
 RT "The genome sequence of *Drosophila melanogaster*.";
 RL Science 287:2185-2195(2000).

RA [2]
 RP SEQUENCE FROM N.A.
 RA Lamsdell S.J., Miller N.S.;
 RL Submitted (JUL-2001) to the EMBL/GenBank/DBJ databases.

RP [3]
 RA SEQUENCE FROM N.A.
 RA Grauso M., Sattelle D.B.;
 RT "Drosophila melanogaster nicotinic acetylcholine receptor betas
 subunit (nAChbeta-21C gene).";
 RL Submitted (JUL-2000) to the EMBL/GenBank/DBJ databases.
 DR EMBL: AE003589: AAF51485.1: -;
 DR EMBL: AJ18761: CAC48166.1: -;
 DR EMBL: AY005148: AAF89090.1: -;
 DR FlyBase: FBgn0031261: nACh-beta-21C.
 DR InterPro: IPR000188: GABA_receptor.
 DR InterPro: IPR001175: Neur_channel.
 DR Pfam: PF02931: Neur_chan_LBD: 1.
 DR Pfam: PF02932: Neur_chan_memb: 1.
 DR PRINTS: PR00252: NRIONCHANNEL.
 KW Receptor.
 SQ
 SSOURCE 441 AA; 49241 MW; A374DBE17791600 CRC64;

Query Match 100.0%; Score 2311; DB 5; Length 441;
 Best Local Similarity 100.0%; Pred. No. 1.6e-203;
 Matches 441; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MTTPTKIKAVYSGPGLPILLQMLMGMLGITSVPGATATADPKANAVKALDRHAGLFT 60
 DB 1 MTTPTKIKAVYSGPGLPILLQMLMGMLGITSVPGATATADPKANAVKALDRHAGLFT 60
 QY 61 NYDSVQPVFGPTNTNSLEKVVYIDIDELNGKLTTHCWLNRDEERYWQPSQYDNI 120
 DB 61 NYDSVQPVFGPTNTNSLEKVVYIDIDELNGKLTTHCWLNRDEERYWQPSQYDNI 120
 QY 121 TQTLKSEVWTPQITLFGNDEGGLMAETQVTLSDGHFRMPRPVAVTAYCELMMLMPH 180
 DB 121 TQTLKSEVWTPQITLFGNDEGGLMAETQVTLSDGHFRMPRPVAVTAYCELMMLMPH 180
 QY 181 DKOSCKLIKISWGLKVVLPENGTAARGESLDHDLVQSPWEIYDSRAHFVSQDYGYMEY 240
 DB 181 DKOSCKLIKISWGLKVVLPENGTAARGESLDHDLVQSPWEIYDSRAHFVSQDYGYMEY 240
 QY 241 TLTQRSSMTAVIYTPASCIYVIALSAFWLPPHMGGEKIMINGLLIYIAAFIMYFAQ 300
 DB 241 TLTQRSSMTAVIYTPASCIYVIALSAFWLPPHMGGEKIMINGLLIYIAAFIMYFAQ 300
 QY 301 LTPVLSNNTPLVYFYSTSLIYSVSTIVELVLYLATGKHKRLPALRKLHGLHGTW 360
 DB 301 LTPVLSNNTPLVYFYSTSLIYSVSTIVELVLYLATGKHKRLPALRKLHGLHGTW 360
 QY 361 LLLSVFTGESQAEKTKEMDEHRYEADDEQESSPLGINHTEYVGAACANQFDMLLATAV 420
 DB 361 LLLSVFTGESQAEKTKEMDEHRYEADDEQESSPLGINHTEYVGAACANQFDMLLATAV 420

QY 421 DRISVFSFLAFLILAIRCSV 441
 DB 421 DRISVFSFLAFLILAIRCSV 441

RESULT 2
 ID Q95R44 PRELIMINARY; PRT; 327 AA.
 AC Q95R44;
 DT 01-DEC-2001 (TREMblrel. 19, Created)
 DT 01-DEC-2001 (TREMblrel. 19, Last sequence update)
 DT 01-DEC-2001 (TREMblrel. 19, Last annotation update)
 DE S009326P.
 GN NACRBETA-21C.
 OS *Drosophila melanogaster* (Fruit fly).
 OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
 OC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
 OC Ephydroidea; Drosophilidae; Drosophila.
 OX NCBI_TaxId=7227;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Stapleton M., Brokstein P., Hong L., Agbayani A., Carlson J.,
 RA Champe M., Chavez C., Dorsett V., Farfan D., Frise E., George R.,
 RA Gonzalez M., Guarin H., Li P., Liao G., Miranda A., Mungall C.J.,
 RA Nunoo J., Pacled J., Paragas V., Park S., Phouanavong S., Wan K.,
 RA Yu C., Lewis S.E., Rubin G.M., Celisner S.;
 RL Submitted (OCT-2001) to the EMBL/GenBank/DBJ databases.
 DR EMBL: AY061622: AAL29170.1: -;
 SQ
 SSOURCE 327 AA; 36457 MW; 104EB50D58779D43 CRC64;

Query Match 71.4%; Score 1650; DB 5; Length 327;
 Best Local Similarity 99.4%; Pred. No. 4.9e-143;
 Matches 310; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 MTTPTKIKAVYSGPGLPILLQMLMGMLGITSVPGATATADPKANAVKALDRHAGLFT 60
 DB 1 MTTPTKIKAVYSGPGLPILLQMLMGMLGITSVPGATATADPKANAVKALDRHAGLFT 60
 QY 61 NYDSVQPVFGPTNTNSLEKVVYIDIDELNGKLTTHCWLNRDEERYWQPSQYDNI 120
 DB 61 NYDSVQPVFGPTNTNSLEKVVYIDIDELNGKLTTHCWLNRDEERYWQPSQYDNI 120
 QY 121 TQTLKSEVWTPQITLFGNDEGGLMAETQVTLSDGHFRMPRPVAVTAYCELMMLMPH 180
 DB 121 TQTLKSEVWTPQITLFGNDEGGLMAETQVTLSDGHFRMPRPVAVTAYCELMMLMPH 180
 QY 181 DKOSCKLIKISWGLKVVLPENGTAARGESLDHDLVQSPWEIYDSRAHFVSQDYGYMEY 240
 DB 181 DKOSCKLIKISWGLKVVLPENGTAARGESLDHDLVQSPWEIYDSRAHFVSQDYGYMEY 240
 QY 241 TLTQRSSMTAVIYTPASCIYVIALSAFWLPPHMGGEKIMINGLLIYIAAFIMYFAQ 300
 DB 241 TLTQRSSMTAVIYTPASCIYVIALSAFWLPPHMGGEKIMINGLLIYIAAFIMYFAQ 300
 QY 301 LTPVLSNNTPLV 312
 DB 301 LTPVLSNNTPLV 312

RESULT 3
 ID Q91X60 PRELIMINARY; PRT; 512 AA.
 AC Q91X60;
 DT 01-DEC-2001 (TREMblrel. 19, Created)
 DT 01-DEC-2001 (TREMblrel. 19, Last sequence update)
 DT 01-DEC-2001 (TREMblrel. 19, Last annotation update)
 DE STIMULAR TO CHOLINERGIC RECEPTOR, NICOTINIC, ALPHA POLYPEPTIDE 2
 DE (NEURONAL).
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 OX NCBI_TaxId=10090;

RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Salivary Gland;
 RA Strausberg R;
 RL Submitted (JUL-2001) to the EMBL/Genbank/DBJ databases.
 DR EMBL: BC011490; AF11490.1; -
 KW Receptor.
 SQ SEQUENCE 512 AA: 58735 MW: AC0C6CC398B89C6C CRC64;

Query Match 18.8%; Score 434; DB 11; Length 512;
 Best Local Similarity 25.0%; Pred. No. 3.7e-31;
 Matches 130; Conservative 91; Mismatches 179; Indels 120; Gaps 19;

OY 9 APVSGPGLPQLQMLMGLMGLTSGVPGATATADPKANANKALDRHAGLFTNYDSYDOP 68
 DB 2 AP-SHAPFQWILYLMCLL-----MPAVLA--OQGSHTHADRLFKHLFGSYNWAR 53
 OY 69 VFGPTPNVSLKVVY-----IDIDELNGKLTTHCMNLNRDEDERVQPSQYDNT 121
 DB 54 V-----PNTSDVYVAFEGSLIAQLIDVDEKNQMTNVMKQEMNDYKLRMDPAEFGNIT 108
 OY 122 QITLKSEVWTPQITLFGNDEGLMAETGYLTH---DGHFRMPRAYVTANCCLMLM 178
 DB 109 SLKVPSEMIIDIVLYNNADGE-FAVTHMTKAHLFTGTVMHVPPIAYKSSCSIDVTF 167
 OY 179 PHDKSCCLKIGSW--GLKVLVPENGTAARGESLDHDLVQSPPEWIVDSRAHFVSODY- 234
 DB 168 PFDQCNCKKFGSMYDRAKIDLEQ---MERTVDLKDWESEMAIINATGYNSKYD 223
 OY 235 -YGYMEYTLTAQRSSMTAVIYTPASCIIVILASFMPLPRHMGKEKIMINGLII 289
 DB 224 CCAEIPDVTYFVIRLPLFTINLIIPCLLISCLTVLFTVLPSC-GEKITLCSVLL 282
 OY 290 VIAAFMTAQLPLVSNTPLVITYSTSLYLSVTEVLEVLYLATGAKHR-----R 344
 DB 283 SLTFVLLLTETIPSTSLVPLIGELTFTMTFTLIVTVFVLNV---HHRSPSTHN 338
 OY 345 LPEALKRLHGHGTWLLS-----VFSTG-----ESQAEKTKEMDEHY 385
 DB 339 MPMWAVALLGKVPRLMNNRPLRPMELHGSPLKLSPTYHMLTMMDEBERETEER 398
 OY 386 EEADEQ-----ESSPLGI-----NHTEVPGAKN----- 409
 DB 399 EEDENICMCAGLPOSSMGLVGHGSLHBRMGPEAKTPSQASEILLSPQIQKALEGVH 458
 OY 410 -----QPDALLATAVDRISVSESLAFLIL 435
 DB 459 IADHLRSEDADSSVKEDMKRYAMVVDRI---FLMLFIIV 494

RESULT 4
 OYX62 PRELIMINARY: PRT; 457 AA.

AC OYX62; PRELIMINARY: PRT; 457 AA.
 DT 01-NOV-1999 (TREMblrel. 12, Created)
 DT 01-NOV-1999 (TREMblrel. 12, last sequence update)
 DT 01-DEC-2001 (TREMblrel. 19, last annotation update)
 DE NICOTINIC ACETYLCHOLINE RECEPTOR ALPHA-SUBUNIT.
 OS Canis familiaris (Dog).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Carnivora; Fissipedia; Canidae; Canis.
 OX NCBI_TaxID=9615;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=MUSCLE;
 RA MEDLINE-20092256; PubMed-10628675;
 RA Yoshioka T., Uzuka Y., Tanabe S., Sarashina T., Ishiguro N.;
 RT "Molecular cloning of the canine nicotinic acetylcholine receptor
 alpha-subunit gene and development of the ELISA method to diagnose
 myasthenia gravis";
 RL Vet. Immunol. Immunopathol. 72:315-324(1999).
 CC -i- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN (BY SIMILARITY).

CC -i- SIMILARITY: BELONGS TO THE LIGAND-GATED IONIC CHANNELS FAMILY.
 DR EMBL: AB021708; BA76876.1; -
 DR InterPro: IPR000188; GABA_receptor.
 DR InterPro: IPR001175; Neur_channel.
 DR Pfam: PF02931; Neur_chan_LBD; 1.
 DR Pfam: PF02932; Neur_chan_memb; 1.
 DR PRINTS: PR00252; NRIONCHANNEL.
 DR PROSITE: PS00236; NEURON_ION_CHANNEL; 1.
 KW Glycoprotein; Ionic channel; Postsynaptic membrane; Transmembrane.
 SQ SEQUENCE 457 AA: 51878 MW: F71AD19AE2ECA3C8 CRC64;

Query Match 18.1%; Score 418.5; DB 6; Length 457;
 Best Local Similarity 26.6%; Pred. No. 8.3e-30;
 Matches 123; Conservative 94; Mismatches 164; Indels 81; Gaps 21;

OY 26 MLMLGTSVPGATATADPKANANKALDRHAGLFTNYDSYDOPV---FQGTPTNVSLEMY 82
 DB 7 LLLGLCSA-GVYVSGSEHT-----RLVAKLFEDYNSVVRVEDHQAVENTVGLQ- 57
 OY 83 VYIIDDELNGKLTTHCMNLNRDEDERVQPSQYDNTQITLKSEVWTPQITLFGNDE 142
 DB 58 IQLINDEVNQIYTTNVRLKQWVDNLRKNPEYGVGVRKIHIPSEKIRMPDLVLYNNAD 117
 OY 143 G-GLMAETQVTLSDHGRMPMPRAYTAVYACELNMLNPHDKSCCLKIGSW---GLKV 196
 DB 118 GDAIYKFKVLVDYTGHTTPRALFKSYCELYVHPPEDQCNCKMKGITWYDOSVYA 177
 OY 197 VLPENGTAARGESLDHDLVQSPPEWIVDSRA-HFVSQDYG-----YME-YTLTAOR 246
 DB 178 INDES-----DQPDLSNFMESEGVVKEKRGKHWV---FYACCPSTPYLDITYHVMOR 229
 OY 247 RSMYNAVITYTPASCIIVILASFMPLPRHMGKEKIMINGLIIVIAAFMTAQLPLVS 306
 DB 230 LPLFTVYVNIIPCLFSEFTGLVFLPTD-SGKMTLSISVLLSLVFLVLYELIPST 288
 OY 307 NNPPLVIVYSTSLYLSVTEVLEVLYLATGAKHR-----LPEALKRLHGHGTWL 361
 DB 289 SAVPLIGKMLFTMVAVIAIITYIVT---NTHHRSTHMPWVKVF---IDTIFP 341
 OY 362 LLSVFSTGESEAOKTKEMDEHY-EEADEQSS-----PLG-----INHTEVPGA--- 406
 DB 342 NIMFSTMKRPSREK---QDKRIFEDIDISDGKPPCPMGHSPFLKHREPKSAIG 398
 OY 407 -----KANQ-----FDWALLATAVDRISVSESLAFLI 434
 DB 399 VYIATMTKSDQESNNAEMKRYAMVMDHILLGVFMVLCII 440

RESULT 5
 OY18C7 PRELIMINARY: PRT; 452 AA.

AC OY18C7; PRELIMINARY: PRT; 452 AA.
 DT 01-OCT-2000 (TREMblrel. 15, Created)
 DT 01-OCT-2000 (TREMblrel. 15, last sequence update)
 DT 01-DEC-2001 (TREMblrel. 19, last annotation update)
 GN ALPHA 10 SUBUNIT OF NICOTINIC ACETYLCHOLINE RECEPTOR PRECURSOR.
 OS Gallus gallus (Chicken).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianine;
 OC Gallus.
 OX NCBI_TaxID=9031;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=SPLEEN;
 RA Barabino B., Boulter J., Alliod C., Ballivet M.;
 RT "Alpha 10 nAChR expression in the developing chick embryo";
 RL Submitted (JUL-2000) to the EMBL/Genbank/DBJ databases.
 CC -i- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN (BY SIMILARITY).
 CC -i- SIMILARITY: BELONGS TO THE LIGAND-GATED IONIC CHANNELS FAMILY.
 DR EMBL: AJ295624; CAB99482.1; -
 DR InterPro: IPR000188; GABAA_receptor.

DR InterPro: IPRO01175; Neur_channel.
 DR Pfam: PF02931; Neur_chan_LBD; 1.
 DR Pfam: PF02932; Neur_chan_memb; 1.
 DR PRINTS: PR00252; NRIONCHANNEL.
 DR PROSITE: PS00236; NEUROTR_ION_CHANNEL; 1.
 KW Glycoprotein; Ionic channel; Postsynaptic membrane; Receptor; Signal;
 FT Transmembrane.
 FT SIGNAL 1 26 POTENTIAL.
 FT CHAIN 27 452 ALPHA 10 SUBUNIT OF NICOTINIC
 FT ACETYLCHOLINE RECEPTOR.
 SO SEQUENCE 452 AA; 50052 MW; 3836C31530BAAD8 CRC64;

Query Match 18.0%; Score 417; DB 13; Length 452;
 Best Local Similarity 28.4%; Pred. No. 1,1e-29;
 Matches 129; Conservative 73; Mismatches 190; Indels 62; Gaps 18;

QY 23 LMGILMGILTSVPGATATADPKNANKALDRHAGLFTNYSDVQPVQF--PTNYSLE 80
 DB 10 LLAISLSTGLVAPCGA-----AQRHAKHLHDLRFANYSALRPA-EDTERALNVTIQ 62
 QY 81 MVTYIDIDELNGKLTTHCWLNRWDERWQPSQYDNTITQITLKSEVWPTQITLF- 138
 DB 63 VTLGQIIDMDERNQVLSYLVNRQAMDLAKMDKDAAGIDSTRIPSSYWRDVIYLN 122
 QY 139 NGDE--GGLMAETQVTLSDGHFRMPRPVYTAACELNMLNMPDKQSKLKIGSWGKLV 196
 DB 123 NADRFEGGSM-ETWVVLRSDDGHIMWDSKAYDVSYPFGQOCRLTFGSM---- 177
 QY 197 VLPRNGTAGESL-----DHDDLVSQPEWEIVDSRA-----HVSQDYYGYMYET 241
 DB 178 -----TYNQIIDLRLNLDLTDGLDIEVENWEVLGMPATRVNVTYGCCEPPDYVT 231
 QY 242 LTAQRSSMYAVIYTPASCIIVILASFWLPPHMGKEKIMINGLLIVIAFLMYAQL 301
 DB 232 LLRRRASFTYFNLLPCVMSFLAPLGFYLPAD-SGKVSGLGYVILALVYFQLVAES 290
 QY 302 LPLVSNTPLVVIFYSTLSLVSTIVEVLVYL-ATGKHRLRLPRLKRLHGLGTM 360
 DB 291 MP-SESVPLIGKYYIATMTITASTALTIFIMNIHCGPGAPRPAPPARLLIHLARA 349
 QY 361 L-LLSVSTGESQAEKTKEM-DEHPYEADQESSP---LGINHTVVP----- 405
 DB 350 LCVEVEGSGCRPREGGMPDPDPEGEVPGICRPSRCLCHHAAVLSVGYTAGVER 409
 QY 406 ---AKANQFDALLATPAVDR-ISVFSFLAFLLI 434
 DB 410 RHRTAQRRAAEMKKVAKYMDRFEMVPEFLMYFLM 443

RESULT 6
 Q9ET51 PRELIMINARY; PRT; 629 AA.
 AC Q9ET51;
 DT 01-MAR-2001 (TREMblrel. 16, Created)
 DT 01-MAR-2001 (TREMblrel. 16, Last sequence update)
 DT 01-DEC-2001 (TREMblrel. 19, Last annotation update)
 DE NICOTINIC ACETYLCHOLINE RECEPTOR ALPHA 4 SUBUNIT.
 GN CHRNA4 OR ACRA4.
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 OC NCBI_TaxID=10090;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=LONG SLEEP SELECTED LINE;
 RA Stitzel J.A., Jimenez M., Smolen A., Modir J.;
 RT "Cloning of mouse nicotinic acetylcholine receptor subunit alpha 4
 RT cDNA";
 RL Submitted (AUG-2000) to the EMBL/GenBank/DBJ databases.
 CC -1- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN (BY SIMILARITY).
 CC -1- SIMILARITY: BELONGS TO THE LIGAND-GATED IONIC CHANNELS FAMILY.
 DR EMBL: AF225912; AAF34716.2; -.

DR MGD; MGI:87888; Chrna4.
 DR InterPro: IPRO01175; Neur_channel.
 DR PRINTS: PR00252; NRIONCHANNEL.
 DR PROSITE: PS00236; NEUROTR_ION_CHANNEL; 1.
 KW Glycoprotein; Ionic channel; Postsynaptic membrane; Receptor;
 FT Transmembrane.
 FT SIGNAL 1 26 POTENTIAL.
 FT CHAIN 27 452 ALPHA 10 SUBUNIT OF NICOTINIC
 FT ACETYLCHOLINE RECEPTOR.
 SO SEQUENCE 452 AA; 50052 MW; 3836C31530BAAD8 CRC64;

Query Match 18.0%; Score 417; DB 11; Length 629;
 Best Local Similarity 28.6%; Pred. No. 1,1e-29;
 Matches 98; Conservative 74; Mismatches 137; Indels 34; Gaps 9;

QY 11 VSGPGLPILQMLMGILTSVPGATATADPKNANKALDRHAGLFTNYSDVQPVF 70
 DB 3 IGGGAPPLPILLLLLLITGLTPASSHETRAHAERLLK---LFGSNKMSRPV- 57
 QY 71 QGTPTNYSLEVVY-----IDIDELNGKLTTHCWLNRWDERWQPSQYDNTITQI 123
 DB 58 ---ANISDVVLVRRGSLIAQLIDVDEKNQMTTNVWVKQEMHDKLRMDPGDVENVTSI 113
 QY 124 TLKSEVWPTQITLFGNDEGGLMAETQVTLSH---DGHFRMPRPVYTAACELNMLNMPH 180
 DB 114 RIPSPELLRPDIYLYNNADGN-FAVTHLTRKHLTYDGVQVTPPAITKSSCIDVTFPPF 172
 QY 181 DKQSKLKIGSW---GLKVLPRNGTAGESLDHDDLVSQPEWEIVDSRAHVSQDY-- 234
 DB 173 DOQCKTMFGSMVYDKAKIDL---VSMHSRVQDLDFWESGEVYIVDAVGTYNRRKECC 228
 QY 235 ---GYMEYITLTAQRSSMTAVIYTPASCIIVILASFWLPPHMGKEKIMINGLLIIV 291
 DB 229 AEIYPTDIYATIRRLPLPYFINILIPCLLISCTVLYFVYSEBC-DEKVLICISVLLST 287
 QY 292 AAFIYFAQLPLVSNTPLVVIFYSTLSLVSTIVEVLVYL 334
 DB 288 TVFLILITEIIPSTSLVPLIGETVLTMTVITVYTVLV 330

RESULT 7
 Q923N8 PRELIMINARY; PRT; 629 AA.
 AC Q923N8;
 DT 01-DEC-2001 (TREMblrel. 19, Created)
 DT 01-DEC-2001 (TREMblrel. 19, Last sequence update)
 DT 01-DEC-2001 (TREMblrel. 19, Last annotation update)
 DE BM40117.2.1 (CHOLINERGIC RECEPTOR, NICOTINIC, ALPHA POLYPEPTIDE
 DE 4 (ISOFORM 1)).
 GN CHRNA4.
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 OC NCBI_TaxID=10090;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Wall M.;
 RL Submitted (APR-2001) to the EMBL/GenBank/DBJ databases.
 DR EMBL: AL450341; CAC40724.1; -.
 KW Receptor.
 SO SEQUENCE 629 AA; 70304 MW; 26145B6ED50B4IC CRC64;

Query Match 17.8%; Score 412; DB 11; Length 629;
 Best Local Similarity 28.6%; Pred. No. 5,2e-29;
 Matches 98; Conservative 73; Mismatches 138; Indels 34; Gaps 9;

QY 11 VSGPGLPILQMLMGILTSVPGATATADPKNANKALDRHAGLFTNYSDVQPVF 70
 DB 3 IGGGAPPLPILLLLLLITGLTPASSHETRAHAERLLK---LFGSNKMSRPV- 57
 QY 71 QGTPTNYSLEVVY-----IDIDELNGKLTTHCWLNRWDERWQPSQYDNTITQI 123
 DB 58 ---ANISDVVLVRRGSLIAQLIDVDEKNQMTTNVWVKQEMHDKLRMDPGDVENVTSI 113

QY 124 TKSSSEVTPQITLFGDEGMAETQVTLSH---DGHRMMPRAVYATACELMIMMMPH 180
 DB 114 RIPSSELMRPDIIVLYNNAGD-FAVTHLTKAHLFYGRQWMPRAIYKSSCIDVFFPF 172
 QY 181 DKOSCKLTGSM---GLKVVLPENGTARGESLDHDDLVOSPEWEIVDSRAHFVSQDY--- 234
 DB 173 DOONCTMKGSMWYDRAKIDL---VSMHSRYDQDLDFWESGEVVIYDAVGTWTRKYECG 228
 QY 235 ---YGMETTLTAQRSSWYTAIVTPASCIYILASAFMLPRHMGCEKIMINGLLIYI 291
 DB 229 AEIYDPITAFIIRLRPLFTYTNLIIPCLLISCLIVLVLPSEC-GEKVTLCTISVLSL 287
 QY 292 AAFLMYFAQLLPVLSNNPPLVVFYSTSLYSVSTIVEVLV 334
 DB 288 TWFLLLITEIIPSTSLVPLIGEYLFTMIFVTLISIVIVFL 330

RESULT 8
 Q923N7 PRELIMINARY: PRT: 641 AA.
 ID Q923N7:
 AC Q923N7:
 DT 01-DEC-2001 (TREMblrel. 19, Created)
 DT 01-DEC-2001 (TREMblrel. 19, Last sequence update)
 DT 01-DEC-2001 (TREMblrel. 19, Last annotation update)
 DE BM401L17.2.2 (CHOLINERGIC RECEPTOR, NICOTINIC, ALPHA POLYPEPTIDE 4 (ISOFORM 2)).
 GN CHRNA4.
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Musinae; Mus.
 OX NCBI_TaxId=10090;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Wall M.;
 RL Submitted (Apr-2001) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AL450341; CAC40725.1; -
 KW Receptor.
 SQ SEQUENCE 641 AA: 71747 MW; BEBPF1164C1C4A0 CRC64;

Query Match 17.8%; Score 412; DB 11; Length 641;
 Best Local Similarity 28.6%; Pred. No. 5.4e-29;
 Matches 98; Conservative 73; Mismatches 138; Indels 34; Gaps 9;

QY 11 VSGPGLPILLQMLMGLMSVPGATATADPKNANKALDRHAGLFTNIDSDVQPVF 70
 DB 3 IGGSGAPPLLPILLPLLLLTGLTGLLPASSHIEFRAHAEERLKR---LFGYKMSRPV- 57
 QY 71 OGTPYVSLSEMYVY-----IDIDELNGKLTTHCMLRLMRDEERYWQSOYDNTQI 123
 DB 58 ----ANISDVAVRFGELSLAQLIDYDERKNOMMTNVMVAQENHDYKLRMDPGYEVNTSI 113
 QY 124 TKSSSEVTPQITLFGDEGMAETQVTLSH---DGHRMMPRAVYATACELMIMMMPH 180
 DB 114 RIPSSELMRPDIIVLYNNAGD-FAVTHLTKAHLFYGRQWMPRAIYKSSCIDVFFPF 172
 QY 181 DKOSCKLTGSM---GLKVVLPENGTARGESLDHDDLVOSPEWEIVDSRAHFVSQDY--- 234
 DB 173 DOONCTMKGSMWYDRAKIDL---VSMHSRYDQDLDFWESGEVVIYDAVGTWTRKYECG 228
 QY 235 ---YGMETTLTAQRSSWYTAIVTPASCIYILASAFMLPRHMGCEKIMINGLLIYI 291
 DB 229 AEIYDPITAFIIRLRPLFTYTNLIIPCLLISCLIVLVLPSEC-GEKVTLCTISVLSL 287
 QY 292 AAFLMYFAQLLPVLSNNPPLVVFYSTSLYSVSTIVEVLV 334
 DB 288 TWFLLLITEIIPSTSLVPLIGEYLFTMIFVTLISIVIVFL 330

RESULT 9
 Q9N508 PRELIMINARY: PRT: 664 AA.
 ID Q9N508
 AC Q9N508;

DT 01-OCT-2000 (TREMblrel. 15, Created)
 DT 01-OCT-2000 (TREMblrel. 15, Last sequence update)
 DT 01-DEC-2001 (TREMblrel. 19, Last annotation update)
 DE HYPOTHETICAL 75.8 KDA PROTEIN.
 GN F27B3.2.
 OS Caenorhabditis elegans.
 OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditidae;
 OC Rhabditidae; Peloderinae; Caenorhabditis.
 OX NCBI_TaxId=6239;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN-BRISTOL N2;
 RX MEDLINE=99069613; PubMed=9851916;
 RA None;
 RT "Genome sequence of the nematode C. elegans: a platform for
 RT investigating biology. The C. elegans Sequencing Consortium.";
 RL Science 282:2012-2018(1998).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN-BRISTOL N2;
 RA Wilson R.;
 RT "The sequence of C. elegans cosmid F27B3.";
 RT Submitted (Mar-1999) to the EMBL/GenBank/DBJ databases.
 RN [3]
 RP SEQUENCE FROM N.A.
 RC STRAIN-BRISTOL N2;
 RA Waterston R.;
 RT "Direct Submission.";
 RT Submitted (SEP-2001) to the EMBL/GenBank/DBJ databases.
 CC -1- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN (BY SIMILARITY).
 CC -1- SIMILARITY: BELONGS TO THE LIGAND-GATED IONIC CHANNELS FAMILY.
 DR EMBL; AC006631; AAF39790.1; -
 DR InterPro; IPR001175; Neur_channel.
 DR PRINTS; PR00252; NRIONCHANNEL.
 DR PROSITE; PS00236; NEURORION_CHANNEL. 1.
 KW Glycoprotein; Hypothetical protein; Ionic channel;
 KW Postsynaptic membrane; Transmembrane.
 SQ SEQUENCE 664 AA: 75804 MW; 636288145947CFAL CRC64;

Query Match 17.8%; Score 411.5; DB 5; Length 664;
 Best Local Similarity 27.3%; Pred. No. 6.3e-29;
 Matches 104; Conservative 81; Mismatches 167; Indels 29; Gaps 10;

QY 27 LMGSLTSP-----GATATADPKNANKALDRHAGLFTNIDSDVQPVF 79
 DB 55 LVGGLEPISVKNKGANGNELSSRADQNM--RLYRDLLDYNNNEVRPSVHSKEPINTVF 111
 QY 80 EMVVT-YIDIDELNGKLTTHCMLRLMRDEERYWQSOYDNTQITLKSSVTPQITLF 138
 DB 112 VESLTIIDYDERNQILTTNSWIRLHWVDYKLVMDRPLYQVTRIRIPSDKIMKPIIILY 171
 QY 139 NDEDEGLM--AETQVTLSHDGHRMMPRAVYATACELMIMMMPHDKOSCKLTGSM--- 192
 DB 172 NNADAQYKMSVMSVDIYVLYGNIHWPLSAIFTSRCLPVKHPFRQCIILKVASWAVD 231
 QY 193 GLKVVLPENGTARGESLDHDDLVOSPEWEIVDSRA-----HFVSQDYGYMEYTTAQR 246
 DB 232 GTRKIDL---LKSSEGDLNTYTNTEWSLIGIRAKKNQVYISCCPEPFIDVHTIR 287
 QY 247 RSSMYTAIVTPASCIYILASAFMLPRHMGCEKIMINGLLIYIAFLMYFAQLPVLVS 306
 DB 288 RAMEFYFNLLIPVLLISLAIMGFYPTD-SGEKVTLGITSLSLSTVFLMAYAEGRPTA 346
 QY 307 NNTPLVIFYSTSLYLSVSTIVEVLVYL-ATGKHRRRLPEALRLKHLGHLGTWLLLSV 365
 DB 347 EALPLIGIVEGYIMVALGTAMTFTVINIHNGVGVYVPRPQLQIFARYLSKILFVRI 406
 QY 366 FSTTGESQAQKTEKDEHPYE 386
 DB 407 EPHYSHHHVRYWYQKEHPTE 427

RESULT 10
Q9BS29 PRELIMINARY; PRT; 484 AA.
AC Q9BS29;
DT 01-JUN-2001 (TReMBLrel. 17, Created)
DT 01-JUN-2001 (TReMBLrel. 17, Last sequence update)
DT 01-DEC-2001 (TReMBLrel. 19, Last annotation update)
DE HYPOTHEICAL 55.9 KDA PROTEIN.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=LONG CARCINOMA;
RA Strausberg R.;
RL Submitted (MAR-2001) to the EMBL/GenBank/DBJ databases.
CC -1- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN (BY SIMILARITY).
CC -1- SIMILARITY: BELONGS TO THE LIGAND-GATED IONIC CHANNELS FAMILY.
DR EMBL; BC004453; AAR04453.1; -
DR InterPro; IPR001105; Myb_DNA_bind.
DR PRINTS; PR001175; Neur_channel.
DR PROSITE; PS00334; MYB_2; UNKNOWN_1.
DR PROSITE; PS00236; NEUROTR_ION_CHANNEL; 1.
KW Glycoprotein; Hypothetical protein; Ionic channel;
KW Postsynaptic membrane; Transmembrane.
SQ SEQUENCE 484 AA; 55894 MW; 74F70567007BCB51 CRC64;

Query Match 17.6%; Score 406; DB 4; Length 484;
Best Local Similarity 26.0%; Pred. No. 1.3e-28;
Matches 125; Conservative 84; Mismatches 201; Indels 70; Gaps 17;

QY 16 LPLLLQMLMGMLGLTSPGATATADPKNAVYKALDRHLHAGLFTNYSQDQVQVFG-CTP 74
DB 5 LAMLLVQOALLALLPPLLAOGAARRSRRTTPALRLSLYLTRKGVPRDMRKP 64
QY 75 TNSLEMYV-TYIDDELNGKLTTHCWLNLWRDERVQSPQYDNTITLKSEVWTP 133
DB 65 TVVSIDVIVAILNVDENQVLTITIMYROYWDEPLQNMEDPDNTIKLSIPDSTIWP 124
QY 134 QITLFNG--DEGLMAETQVTLSDHGFRMPRAVYTAACGLNMLNMPHDKOSCKLKIGS 191
DB 125 DI-LINEFVVDGKSPNIPYVIRHQGEVQNYKPLQVYACSLDITDYNPEFQVQNSLTFTS 183
QY 192 MGLKVV-----LPENGARGESLDHDDLVQSPMEIYDSRAHF-----VSQDYYGY 237
DB 184 W-LHTIDINISLRLEBKYS-----DRSVFMNGEWEELGVLPYRFESMESSNYAE 237
QY 238 MEYTLTAORRSMYTAIVYTPASCIYIALSAFWLPPHMGGEKIMINGLIITVIAAFMY 297
DB 238 MKFYVIRRRPLFYVSSLLPSIFLWMDIVGYLPPN-SESRVSFKITLLGYSVFLII 296
QY 298 FAQLLPVLSNTPLVIFYSISLYSVTIVEVLVLYLAKGKKRL-----PALRKLL 353
DB 297 VSDLPATAGTPLIGYFVVCMLLVISLAETIFIVRLV---HKODIQDPVPAMLRLV 353
QY 354 HGLHGLTLLSVSTGE-----SOAEKTK-----MDEH-----PYEADDESSP- 395
DB 354 LERT-ANLLCLIREGSTSORPATKTDSCSMGNHCSMGSPRODFEKKPRRCSPFP 412
QY 396 -----LGINTEVVGAKANQF-----DMLLATAVDRISFVSFLAFLIAT 437
DB 413 PPREASLAVCGLLQELSSIRQFLERKDEIRVARDWLRVGSVLDKLLFHYILLAVLSI 472

RESULT 11
Q9JUB5 PRELIMINARY; PRT; 447 AA.
AC Q9JUB5;
DT 01-OCT-2000 (TReMBLrel. 15, Created)
DT 01-OCT-2000 (TReMBLrel. 15, Last sequence update)

DT 01-DEC-2001 (TReMBLrel. 19, Last annotation update)
DE NEURONAL NICOTINIC ACETYLCHOLINE RECEPTOR SUBUNIT.
GN ALPHA10.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=SPRAGUE-DAWLEY;
RX MEDLINE=21145873; PubMed=11248107;
RA Elgoyhen A.B., Vetter D.E., Katz E., Rothlin C.V., Heinemann S.F.,
RA Boulter J.;
RT "alpha 10: A determinant of nicotinic cholinergic receptor function in
RT mammalian vestibular and cochlear mechanosensory hair cells."
RL Proc. Natl. Acad. Sci. U.S.A. 98:3501-3506(2001).
CC -1- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN (BY SIMILARITY).
CC -1- SIMILARITY: BELONGS TO THE LIGAND-GATED IONIC CHANNELS FAMILY.
DR EMBL; AF196344; AAR27624.1; -
DR InterPro; IPR001175; Neur_channel.
DR PRINTS; PR00252; NRIONCHANNEL.
DR PROSITE; PS00236; NEUROTR_ION_CHANNEL; 1.
KW Glycoprotein; Ionic channel; Postsynaptic membrane; Receptor;
KW Transmembrane.
SQ SEQUENCE 447 AA; 49819 MW; EEE49D93490B698F CRC64;

Query Match 17.5%; Score 404.5; DB 11; Length 447;
Best Local Similarity 27.7%; Pred. No. 1.5e-28;
Matches 125; Conservative 78; Mismatches 191; Indels 57; Gaps 17;

QY 24 MGMLMGITSPGATATADPKNAVYKALDRHLHAGLFTNYSQDQVQVQVQVPT-NVSEMY 82
DB 10 LGFLL--LFLPACLGAEGRLL-----KLFRLFANYTSALRPVADDTQTLNVTLEV 62
QY 83 VT-YIIDDELNGKLTTHCWLNLWRDERVQSPQYDNTITLKSEVWTPQTLTN-- 139
DB 63 LSQITIDDERQVLTLYLMIHQEWTDAYLHMDPRAYGDLAIRPSRLVWMPDITVLYNKA 122
QY 140 GDEGLMAETQVTLSDHGFRMPRAVYTAACGLNMLNMPHDKOSCKLKIGS--GLKV 196
DB 123 DTQPPASSTNVVVRHDAVMDAPAITRSSCRDVSAPFPDADRCGLTESWTHGHOL 182
QY 197 -VLEPENGARGESLDHDDLVQSPMEIYDSRA-----HFVSQDYGYMEYTLTAORRS 249
DB 183 DVPR-----RGTSASLADVEVNEVRVYLGMPARRRVLTYYGCCSPYDPVPTLLRRRA 237
QY 250 MYTAIVTPASCIYIALSAFWLPPHMGGEKIMINGLIITVIAFMYFAQLLPVLSNNT 309
DB 238 AYVCNLLPCVFISLPLAFLPAD--SGEKVSGVTVLLALTVFOLILAESMP-RESV 295
QY 310 PLVVTFYSTSLVSVTIVEVLVLYL-ATGKHRRRLPEALRKLLHGLTWLLSVEST 368
DB 296 PLIGKYMATMTVTFPSALITILMLNHGCPNHPAPARAVKRVLLGLHAKGLCV---RE 352
QY 369 TGESQAEKTKEMDHPYEADDESSPLG-----INHTEV-----PGA 406
DB 353 RGE-PCGSKPLESAP--SLQPPASAPGCPHEBRCLCHQALHLHIASTAFRSHRA 409
QY 407 KANQDMALLATAVDRISFVSFLAFLIAT 437
DB 410 QRHEDWKRRLAVMDREFLGIFFCMALVMSL 440

RESULT 12
Q9ROW9 PRELIMINARY; PRT; 494 AA.
AC Q9ROW9;
DT 01-MAY-2000 (TReMBLrel. 13, Created)
DT 01-MAY-2000 (TReMBLrel. 13, Last sequence update)
DT 01-DEC-2001 (TReMBLrel. 19, Last annotation update)
DE NICOTINIC ACETYLCHOLINE RECEPTOR SUBUNIT ALPHA6.
GN CHRNA6 OR NTCAG.

OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 OX NCBI_TaxID=10090;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Le Novere N.;
 RL Submitted (AUG-1999) to the EMBL/GenBank/DBJ databases.
 RN [2]
 RP SEQUENCE FROM N.A.
 RA Marcubio L.M., Champiaux N., Changeux J.P.;
 RT "Cloning of the nicotinic acetylcholine receptor subunit alpha6 from mouse brain."
 RL Submitted (AUG-1999) to the EMBL/GenBank/DBJ databases.
 CC -1- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN (BY SIMILARITY).
 CC -1- SIMILARITY: BELONGS TO THE LIGAND-GATED IONIC CHANNELS FAMILY.
 DR EMBL; AJ243706; CAB3472.1; -.
 DR MGD; MGI:106213; Chrna6.
 DR InterPro; IPR001188; GABAA_receptor.
 DR InterPro; IPR001175; Neur_channel.
 DR Pfam; PF02933; Neur_chan_LBD.1.
 DR Pfam; PF02932; Neur_chan_memb.1.
 DR PRINTS; PR00252; NRIONCHANNEL.
 DR PROSITE; PS00236; NEUROTR_ION_CHANNEL.1.
 DR GlycoProtein; Ionic channel; Postsynaptic membrane; Receptor;
 KW Transmembrane.
 SQ SEQUENCE 494 AA; 56807 MW; 650E6CD5CFCD745F CRC64;

Query Match 17.2%; Score 398; DB 11; Length 494;
 Best Local Similarity 24.3%; Pred. No. 7,1e-28;
 Matches 112; Conservative 84; Mismatches 172; Indels 92; Gaps 15;

QY 52 DRLAGLTNTNDSVQPPFQ-GTPNNVSLQVYTI-DIDELNGKLTHCMLNLRQDE 109
 DB 35 EQLFRLPAHNRFLRPENVSDEPTVHLELAITQLANVEDVNOIMETNMLRHIMKDYR 94
 QY 110 RVMQSOYDNTITQTLKSEVWTPQITLFGDEGLMAE--TOVTLSDHGFRMPAY 167
 DB 95 LRMDPTEIDGIEFLVNPADNITKRPDIYLYNNAVGPQVEGKTKALKIDGVTWTPPAIF 154
 QY 168 TAYCELNLNMPHDKOSCKLIGSW-----GLKVLPENGTANGESLDHDLVQSPDM 220
 DB 155 KSSCPMDITFPFQHONCSLKGFSWTYDKAEIDLII-----GSVYDNMDPFENSEM 206
 QY 221 EIVDSRA--HYVS---QDYGYMEYITLTAQRSSMTAVITTPASCIYIILASFWLPP 274
 DB 207 EIVDSAGYKHDIKYNCEIYTDIYSFYIRLPMFYINILIPCLFISFLVLFVLPSP 266
 QY 275 HNGEKIMINGLIIVIAFLYFAQLPVLNNPGLVIVFYSTLSLVSSTIVEVYL 334
 DB 267 DC-GEKVLICSVLSLVFLVLTETIPSTSLVPLVGEYILFTWIVTSLVTVTVL 325
 QY 335 ---YLATGKHRRRLPEALRKLLHGLGTWLLS-----VF 366
 DB 326 NIHYRTPAH--TMPKWKITIFLQAFPSILMKRPLDKTEAGVKDPSHTKRPAYKF 383
 QY 367 STTGSQ-----AEKTKEMDEPYEADQESSPIGT--- 398
 DB 384 THRGESKLLKBCNHCQSSDIAPGKRSSQOPARVAVANSEHSDVEVIESVQPIAENM 443
 QY 399 -NHFEVPGAKANQFDWALLATAYRISFVSFLAFLIAI 437
 DB 444 KSHNE---TNEVEDDKYIMAVVNV---FLWVETIYCV 476

RESULT 13
 Q964N6 ID PRELIMINARY: PRT: 502 AA.
 AC Q964N6
 DT 01-DEC-2001 (TREMBLrel. 19, Created)
 DT 01-DEC-2001 (TREMBLrel. 19, Last sequence update)
 DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)

DE NICOTINIC ACETYLCHOLINE RECEPTOR ALPHA SUBUNIT.
 GN UNC-63.
 OS Caenorhabditis elegans.
 OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditidae;
 OC Rhabditidae; Peloderinae; Caenorhabditis.
 OX NCBI_TaxID=6239;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Culetto E., Baylis H.A., Morgan N.P., Matsuda K., Fleming J.T.,
 RA Squire M.D., Lewis J.A., Sattelle D.B.;
 RT "The C. elegans unc-63 gene encodes a novel a subunit able to form a functional heteromeric nicotinic acetylcholine receptor when co-expressed in Xenopus oocytes with the products of lev-1 and unc-29."
 RL Submitted (JUL-2000) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AF288374; AAK3056.1; -.
 KW Receptor.
 SQ SEQUENCE 502 AA; 57391 MW; 3C09596F529DE6CB CRC64;

Query Match 17.2%; Score 397.5; DB 5; Length 502;
 Best Local Similarity 23.4%; Pred. No. 8e-28;
 Matches 112; Conservative 97; Mismatches 162; Indels 107; Gaps 16;

QY 43 PKNAVKALDRLAGLTNYSVQPPFQGTPTNVSLEWVY-----IDIDELNGKL 95
 DB 20 PTHANRDA-NRFLFDLADYNKLVAPSENGE-----LVYTFKIKLSQLLDVHEKNQIM 73
 QY 96 TTHCWLNLNRDDEKRWQSDYDNTITQTLKSEVWTPQITLFGDEGLMAE--TOVTL 153
 DB 74 TTNVWLQHSMDYKLRMPDVEGVEVLYVPSDTIMLPDVLVYNADNGYQVTIMTKAKL 133
 QY 154 SHDGHFRMPAYVAYCELNLMNPHDKOSCKLIGSW---GLKVLPENGTANGESLD 210
 DB 134 TTNQVEWAPATYKSMQIDVEFPDQCEKMGFGSWTYGGLVDLQHRDKHLEKTE 193
 QY 211 HD-----DLVQSPMEIYD-----SRAHFYSDYGYMEYTL 242
 DB 194 EDVEGVDPGTEKIVWVVRGIDLSDYPSVENDILNVGKRKSKYRPPCESEFIDITEI 253
 QY 243 TAQRSSMTAVITTPASCIYIILASFWLPRHNGEKIMINGLIIVIAFLYFAQL 302
 DB 254 HLRKRTLFYTVNLIPPSVIGSIFLTLVFLPSD-GEFISLISLISLISLTVFLLVETI 312
 QY 303 PLSNNTPLVIVFYSTLSLVSSTIVEVYL---YLATGKHRRRLPEALRKLLHGLGT 359
 DB 313 PSTSLVITPLIGKYLFTWLVTVLSVVTYVTLNVHRSPTTH--TMPKWKRLPEVDFLPK 370
 QY 360 WLLSVFSTTGSQAETKEMDEHPYEADQESS-PLGINH-----TEV--PGAKANQ- 410
 DB 371 YLDM---TRPQPCGHNSK-----PNRKFDSASRFSIGVNVLVQNSLSPGLNSURE 421
 QY 411 -----EDWALLATAYDRISVSFLA 431
 DB 422 ESSFTLPRDNSFVSVAIYADHLKNEEDDKQVIEDMKYISVMDRIPLITPFA 479

RESULT 14
 ID PRELIMINARY: PRT: 523 AA.
 AC Q96128
 DT 01-JUN-1998 (TREMBLrel. 06, Created)
 DT 01-JUN-1998 (TREMBLrel. 06, Last sequence update)
 DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)
 DE NICOTINIC ACETYLCHOLINE RECEPTOR ALPHA1 SUBUNIT PRECURSOR.
 GN NACHR.
 OS Heliothis virescens (Noctuid moth) (Owllet moth).
 OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
 OC Pterygota; Neoptera; Endopterygota; Lepidoptera; Glossata; Ditrysia;
 OC Noctuidae; Noctuidae; Heliothinae; Heliothis.
 OX NCBI_TaxID=7102;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Jafarigorzini S., Maelicke A.;

"Preliminary: Cloning of nicotinic acetylcholine receptor subunits of *Rt* *Heliothis virescens*.";
 RT Submitted (JUL-1997) to the EMBL/GenBank/DBJ databases.
 CC - SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN (BY SIMILARITY).
 CC - SIMILARITY: BELONGS TO THE LIGAND-GATED IONIC CHANNELS FAMILY.
 DR EMBL: AJ000399; CAN04056.1; -.
 DR InterPro: IPR001188; GABAA_receptor.
 DR InterPro: IPR001175; Neur_channel.
 DR Pfam: PF02931; Neur_chan_LBD; 1.
 DR Pfam: PF02932; Neur_chan_memb; 1.
 DR PRINTS: PR00252; NRIONCHANNEL.
 DR PROSITE: PS00236; NEUROTR_ION_CHANNEL; 1.
 KW Glycoprotein; Ionic channel; Postsynaptic membrane; Receptor; Signal;
 KW Transmembrane.
 FT SIGNAL 1 27 POTENTIAL.
 FT CHAIN 28 523 NICOTINIC ACETYLCHOLINE RECEPTOR.
 SQ SEQUENCE 523 AA; 59214 MW; 7E39A035B6BA00FD CRC64;

Query Match 17.1%; Score 395; DB 5; Length 523;
 Best Local Similarity 23.7%; Pred. No. 1.4e-27;
 Matches 115; Conservative 91; Mismatches 188; Indels 92; Gaps 13;

QY 26 MLMLGTSVGGATATAPKNNANVALDRLHAGLFTNDSYQVPFGCTPTVSLDMVY-- 83
 DB 14 LLVLAALAGCAAMPK-----RLYDDLSTNYKLVLPV-----NVSALTVRI 60
 QY 84 ----YIDIDELNGKLTTHCMLNLRWDEERYWQPSQYDNIQTITLKSSVMTPTLF 138
 DB 61 KLIKSLQIDLVNKLQMTNTNLMVQSWDYKLSWEPREYGVEMLHVPDSHIMRPDIYLY 120
 QY 139 NGDEGG--LMAEQVTLSDHGFRRMPRAYTAYCELMNLMPHDKOSCKLKIGSW--G 193
 DB 121 NMDGNEFVTLATKATINTYGRVEMRPATYKSCSIDVEYFPDQCTCVMKFGSWTYDG 180
 QY 194 LKVVLPENGARGES-----LDHDDLYQSPMEIYD-----SRHHVSQDYGYGMEYTL 242
 DB 181 FQVDLRHIDERGTVNELGVDLSEFYTSVMDLLEVPANNEKFFYCCDDPYLDITFNI 240
 QY 243 TAQRSSMYTAVIYTPASCIYIALSAFWLPPHMGKEKIMINGLLIYIAFLMYFAQL 302
 DB 241 TMRKTFEYTNLITPCMGISFLVLYFPLSD--SGEVSLSITLSLTYFPLLEAIT 299
 QY 303 PVSNNTPV--VIFYSTSLYLSVSTIVEVLYLYLATGKHRRLEALRRLHGLHGTW 360
 DB 300 PPTSLVPLDLCKEFLFMILDTFESICVTVVVLANVHFRSPQHTMAWVRVFIH-VLPR 358
 QY 361 LLL-----SVSTGESQAETKEMDEHP----- 384
 DB 359 LVMRRPHYRVDPHRSRFGALVTAVSESAPWEDGSPLAGLAGGPGPPACACRSCRLH 418
 QY 385 -----YEADQESSPLGINHTEVPGAKANOF-----DMALLATAVDRISFVS 428
 DB 419 DAPALCALRMHRCPELNAIDGINIYAEQTRKEESTRVEDMKYVAMLDLFLMITF 478
 QY 429 SLAFLI 434
 DB 479 TLAVVV 484

RESULT 15
 Q96FB8 PRELIMINARY; PRT; 501 AA.
 AC Q96FB8;
 DT 01-DEC-2001 (TREMblrel. 19, Created)
 DT 01-DEC-2001 (TREMblrel. 19, Last sequence update)
 DE 01-DEC-2001 (TREMblrel. 19, Last annotation update)
 DE SIMILAR TO CHOLINERGIC RECEPTOR, NICOTINIC, BETA POLYPEPTIDE 1 (MUSCLE).
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 OX NCBI_TaxID=9606;

RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=EYE, AND RETINOBLASTOMA;
 RA Strausberg R.;
 RT Submitted (JUL-2001) to the EMBL/GenBank/DBJ databases.
 DR EMBL: BC011371; AAH11371.1; -.
 KW Receptor.
 SQ SEQUENCE 501 AA; 56698 MW; 365CBFA795A51394 CRC64;

Query Match 17.0%; Score 392.5; DB 4; Length 501;
 Best Local Similarity 22.6%; Pred. No. 2.3e-27;
 Matches 117; Conservative 91; Mismatches 178; Indels 131; Gaps 17;

QY 17 PLLIOMLMGLMGLTSVGGATATADPKNNANVALDRLHAGLFTNDSYQVPFGCTPT 75
 DB 3 PGALLMLGLL--GAPLAPGVRS-----EADGRRLRELFEGSYSSVAPAREVDNRV 52
 QY 76 NVSLDMVYT-YIDIDELNGKLTTHCMLNLRWDEERYWQPSQYDNIQTITLKSSVMTPTQ 134
 DB 53 RVSGLIQAQISINKEDEEMSTKRYLDLEWTDYRLSWDPAHGDIDSLRTAESVWLPD 112
 QY 135 ITLNGDEGG--LMAEQVTLSDHGFRRMPRAYTAYCELMNLMPHDKOSCKLKIGSW 192
 DB 113 VLLNNMDGNEFDVALDISVAVSSDGSVWQPGIYRSSCSIQVTFPFDWQNCITMVESSY 172
 QY 193 -----GLKVVLPENGARGESLDHD-DLYQSPMEIYDSDRAHFVSGO--DYGYME-- 239
 DB 173 SYDSSEVSLQYGLGPDQGHQEHHEGTFLENQOWEILHFRSLIQPRGPRGREGQR 232
 QY 240 ----YTLAQRSSMYTAVIYTPASCIYIALSAFWLPPHMGKEKIMINGLLIYIAFL 295
 DB 233 QEVIFYLIRRKPLFYLVNVIAPCILITLLAIFEVYLPD--AGEKMGISIFALLTLTVFL 291
 QY 296 MYFAQLLPVLSNNTPVLYEYFSTSLYLSVSTIVEVLYLY-ATGKHRRLEALRRLH 354
 DB 292 LLLADKVPETSLVPLITIKYLMFTVLTVEFSVLLVAVLNHRSPTTHQPLWVRQIF 351
 QY 355 GHIGTLLLSVSTGESQAETKEMDEHPYEADQESSPLGINHTEVPGA----- 406
 DB 352 HKLPLVRL-----KRPRERDLMEPP-----HSSPSGSGRGRTDE 389
 QY 407 -----KANOF----- 411
 DB 390 YFIRKPPSDFLPKPNRFOPELASAPDLRRTIDGNRAVALPELREVVSSISYAROLQ 449
 QY 412 -----DMALLATAVDRI--SPVSF--SLAFLIL 435
 DB 450 QEDHDALKEWQFVAMVVDRLFLMTFTIFTSVGTLYI 486

Search completed: July 18, 2002, 10:29:21
 Job time: 669 sec

Fri Jul 19 08:26:24 2002

us-09-732-680a-2.rspt

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: July 18, 2002, 10:16:52 ; Search time 35.01 Seconds

(without alignments)
487.727 Million cell updates/sec

Title: US-09-732-680A-2
Perfect score: 2311
Sequence: 1 MTTTPKIKAPVSGPLDLL.....RISFVSFLAFLITAIRCSV 441

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 105224 seqs, 38719550 residues
Total number of hits satisfying chosen parameters: 105224

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : SwissProt_40:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	444.5	19.2	528	1	ACH2_CHICK
2	427	18.5	457	1	ACHA_BOVIN
3	423	18.3	511	1	ACH2_RAT
4	418	18.1	627	1	ACH4_HUMAN
5	416	18.0	529	1	ACH2_HUMAN
6	414.5	17.9	457	1	ACHA_RAT
7	413.5	17.9	493	1	ACH6_RAT
8	412.5	17.8	455	1	ACHA_MOUSE
9	410	17.7	450	1	ACHO_CHICK
10	409	17.7	622	1	ACHX_HUMAN
11	407.5	17.6	557	1	ACH4_CHICK
12	407	17.6	557	1	ACH1_SCHER
13	407	17.6	630	1	ACH4_RAT
14	404.5	17.5	478	1	5HT3_HUMAN
15	403.5	17.5	519	1	ACHO_RAT
16	402.5	17.4	519	1	ACH4_DROME
17	400.5	17.3	457	1	ACH2_XENLA
18	398.5	17.2	457	1	ACH1_XENLA
19	397.5	17.2	468	1	ACH5_HUMAN
20	397	17.2	479	1	ACH9_RAT
21	397	17.2	505	1	ACHB_BOVIN
22	394	17.0	494	1	ACH6_HUMAN
23	393.5	17.0	479	1	ACH9_HUMAN
24	392	17.0	494	1	ACH6_CHICK
25	391.5	16.9	512	1	ACH3_CARAU
26	391	16.9	458	1	ACHO_HUMAN
27	391	16.9	493	1	ACH2_CAEL
28	390.5	16.9	456	1	ACHA_BRARE
29	390.5	16.9	456	1	ACHA_CHICK
30	390.5	16.9	507	1	ACH7_CAEL
31	388	16.8	482	1	ACHA_HUMAN
32	386.5	16.7	461	1	ACHA_TORMA
33	386.5	16.7	499	1	ACH3_RAT

ALIGNMENTS

RESULT ID	ACH2_CHICK	STANDARD:	PRT:	528 AA.
AC	P09480;			
DT	01-MAR-1989 (Rel. 10, Created)			
DT	01-MAR-1989 (Rel. 10, Last sequence update)			
DT	16-OCT-2001 (Rel. 40, Last annotation update)			
DE	Neuronal acetylcholine receptor protein, alpha-2 chain precursor.			
OS	Gallus gallus (Chicken).			
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
OC	Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;			
OC	Gallus.			
OX	NCBI_TaxID=9031;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RC	Tissue=Brain;			
RX	MEDLINE=88283624; PubMed=3267226;			
RA	Nef P., Oneyser C., Allod C., Couturier S., Ballivet M.;			
RT	"Genes expressed in the brain define three distinct neuronal			
RT	nicotinic acetylcholine receptors.";			
RL	EMBO J. 7:595-601(1988).			
CC	- FUNCTION: AFTER BINDING ACETYLCHOLINE, THE ACHR RESPONDS BY AN			
CC	EXTENSIVE CHANGE IN CONFORMATION THAT AFFECTS ALL SUBUNITS AND			
CC	LEADS TO OPENING OF AN ION-CONDUCTING CHANNEL ACROSS THE PLASMA			
CC	MEMBRANE.			
CC	- SUBUNIT: NEURONAL ACHR SEEMS TO BE COMPOSED OF TWO DIFFERENT			
CC	TYPE OF SUBUNITS: ALPHA AND NON-ALPHA (ALSO CALLED BETA). A			
CC	FUNCTIONAL RECEPTOR SEEMS TO CONSIST OF TWO ALPHA-CHAINS AND			
CC	THREE NON-ALPHA CHAINS.			
CC	- SUBCELLULAR LOCATION: Integral membrane protein.			
CC	- SIMILARITY: BELONGS TO THE LIGAND-GATED IONIC CHANNELS FAMILY.			
CC	-----			
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration			
CC	between the Swiss Institute of Bioinformatics and the EMBL Outstation -			
CC	the European Bioinformatics Institute. There are no restrictions on its			
CC	use by non-profit institutions as long as its content is in no way			
CC	modified and this statement is not removed. Usage by and for commercial			
CC	entities requires a license agreement (see http://www.isb-sib.ch/announce/			
CC	or send an email to license@sib-sib.ch).			
CC	-----			
CC	EMBL: X07339; CAB59645.1; JOINED.			P09690 mus musculus
CC	EMBL: X07340; CAB59645.1; JOINED.			P11230 homo sapien
CC	EMBL: X07341; CAB59645.1; JOINED.			O93149 caenorhabdi
CC	EMBL: X07342; CAB59645.1; JOINED.			P02710 torpedo cal
CC	EMBL: X07343; CAB59645.1; JOINED.			P26152 gallus gall
CC	EMBL: X07344; CAB59645.1; JOINED.			P49582 carassius a
CC	EMBL: X07345; CAB59645.1; JOINED.			P13908 mus musculu
CC	EMBL: AJ250360; CAB59625.1; JOINED.			P13908 carassius a
CC	PIR: S00377; ACCH2N.			P36544 homo sapien
CC	InterPro: IPR000188; GABAA_receptor.			P25109 rattus norv
CC	InterPro: IPR001175; Neur_channel.			P02712 torpedo cal
CC	Pfam: PF02931; Neur_chan_LBD; 1.			P02713 gallus gall
CC	Pfam: PF02932; Neur_chan_membr; 1.			
CC	PRINTS: PR00252; NRIONCHANNEL.			
CC	PROSITE: PS00236; NEUROTRN_ION_CHANNEL; 1.			
CC	Postsynaptic membrane; Ionic channel; Glycoprotein; signal;			
CC	Transmembrane; Multigene family.			

```

FT SIGNAL 1 23
FT CHAIN 24 528 NEURONAL ACETYLCHOLINE RECEPTOR PROTEIN,
FT DOMAIN 24 239 ALPHA-2 CHAIN.
FT TRANSMEM 240 264 EXTRACELLULAR.
FT TRANSMEM 272 290
FT TRANSMEM 306 327
FT DOMAIN 328 327
FT TRANSMEM 502 520 CYTOPLASMIC.
FT DOMAIN 394 410
FT DISULFID 158 172 POLY-GLU.
FT DISULFID 222 223 BY SIMILARITY.
FT CARBOHYD 54 54 ASSOCIATED WITH RECEPTOR ACTIVATION
FT CARBOHYD 104 104 (BY SIMILARITY).
SQ SEQUENCE 528 AA; 60675 MW; E76C6360AF876364 CRC64; N-LINKED (GLCNAC. . .) (POTENTIAL).

```

Query Match 19.28; Score 444.5; DB 1; Length 528;
 Best Local Similarity 27.7%; Pred. No. 3.1e-28;
 Matches 119; Conservative 82; Mismatches 147; Indels 81; Gaps 15;

```

QY 39 ATADPRKANAKALDLRLAGLFTNVDSDVOPFQSTPTNVSLDMVTVY-----IDDEL 91
DB 22 ATRKQKOPHGFADRLFKHLFTGYNRSRP-----PNTSDVIVKFGSLAQILDVDEK 76
QY 92 NGKLTTHCWLINLRWDEERWQPSQYDNITQITLKSEVWTPQITLFGDEGLMAETQV 151
DB 77 NQMTTWVWLKQEWSDYKLRNRPEDFQNVTSIRVPSIMWIPDLVLYNNADGE-FAYTHM 135
QY 152 TLSH---DGEFRMMPVAVYATACELNMLNMPHDKSKCKLIGSV---GLKVVLPENGSTAR 205
DB 136 TKHLFENGKVKWPPAYVSSCSIDVTYFPFQDQCKMKFGSWTYDKAKIDL----- 188
QY 206 GESLDL---DDLQSPSEWELVDSRAHFVSQDY-----GYMYTTLAQGRSSMTAVI 255
DB 189 -ENNEHHVDLKDWESEGMATINAGRYNSKRYDCTEIRPDITFYVIRLPLEYITNL 247
QY 256 YTPASCIIVIALSAFMPHPHMGKIKIMINGLIIVIAFLMFAQLPVLNNPPLVAVIF 315
DB 248 IIPOLLISCLVLFVFLPSPC-GEKITLCLSVSLTLVFLILITIPISVLPVLPGEY 306
QY 316 YTSLSLYLSTVIVEVLVLYLANGKHKR-----RLPALRKLHLGLGW-----LLL 363
DB 307 LLEFTMIFVTLIIITVEVLAVY---HHRSPSTHMPHNVSEFGFIPRMLEMKRPLLL 362
QY 364 SVFSTTG-----ESQAEKTKEMDEHYERADQESSPLGINHTEY 403
DB 363 PASGTTQYDPPTGRLSTSRQWLETDVDQKWEDEEEEEEKAKVP-----SRV 417
QY 404 P--GAKANQ 410
DB 418 PSQSGSQCTQ 426

```

RESULT 2
 ACHA_BOVIN STANDARD; PRT; 457 AA.
 AC P02709;
 DT 21-JUL-1986 (Rel. 01, Created)
 DT 21-JUL-1986 (Rel. 01, Last sequence update)
 DT 01-OCT-1996 (Rel. 34, Last annotation update)
 DE Acetylcholine receptor protein, alpha chain precursor.
 GN CHRNA1.
 OS Bos taurus (Bovine).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidea;
 OC Bovidae; Bovinae; Bos.
 ON NCBI_TaxID=9913;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA MEDLINE=84039794; PubMed=6688857;
 RA Noda M., Furutani Y., Takahashi H., Toyosato M., Tanabe T.,

```

RA Shimizu S., Kiyotani S., Kayano T., Hirose T., Inayama S., Numa S.;
RT "Cloning and sequence analysis of calf cDNA and human genomic DNA
RT encoding alpha-subunit precursor of muscle acetylcholine receptor.";
RL Nature 305:818-823(1983).
CC -!- FUNCTION: AFTER BINDING ACETYLCHOLINE, THE ACHR RESPONDS BY AN
CC EXTENSIVE CHANGE IN CONFORMATION THAT AFFECTS ALL SUBUNITS AND
CC LEADS TO OPENING OF AN ION-CONDUCTING CHANNEL ACROSS THE PLASMA
CC MEMBRANE.
CC -!- SUBUNIT: PENTAMER OF TWO ALPHA CHAINS, AND ONE EACH OF THE BETA,
CC DELTA, AND GAMMA (IN IMMATURE MUSCLE) OR EPSILON (IN MATURE
CC MUSCLE) CHAINS.
CC -!- SUBCELLULAR LOCATION: Integral membrane protein.
CC -!- SIMILARITY: BELONGS TO THE LIGAND-GATED IONIC CHANNELS FAMILY.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC EMBL: X02509; CA26345.1; -.
DR PIR: A03169; ACBOAL.
DR InterPro: IPR000188; GABAA_receptor.
DR Pfam: IPR001175; Neur_channel.
DR Pfam: PF02931; Neur_chan_LBD. 1.
DR Pfam: PF02932; Neur_chan_membr. 1.
DR PRINTS: PR00252; NRIONCHANNEL.
DR PROSITE: PS00236; NEUROTR_ION_CHANNEL. 1.
KW Receptor; Postsynaptic membrane; Ionic channel; Glycoprotein; Signal;
KW Transmembrane.
FT SIGNAL 1 20
FT CHAIN 21 457
FT DOMAIN 21 230 ACETYLCHOLINE RECEPTOR PROTEIN, ALPHA
FT TRANSMEM 231 255 CHAIN.
FT TRANSMEM 263 281 EXTRACELLULAR.
FT TRANSMEM 297 316
FT DOMAIN 317 428
FT TRANSMEM 429 447
FT DISULFID 148 162
FT DISULFID 212 213
FT CARBOHYD 161 161
SQ SEQUENCE 457 AA; 51947 MW; 40261D335FD47A32 CRC64;

```

Query Match 18.58; Score 427; DB 1; Length 457;
 Best Local Similarity 27.4%; Pred. No. 6.5e-27;
 Matches 129; Conservative 91; Mismatches 163; Indels 88; Gaps 22;

```

QY 17 PLLQMLMGMLNGLTSPGATADPRKANVKAIDLRLGFTNVDSDVQPY---FQGT 73
DB 5 PLLL--TLGLCSAGL--VLGSEHET-----RLVAKLFEDINSVRPVEDHRAQV 49
QY 74 PTNVSLDMVTVYIDIDELNKLITTCWLNLRWDEERWQPSQYDNITQITLKSEVWTP 133
DB 50 EVTVGLQL-IQLINDEVNQIVTNVRLKQQWVDYNNKMPDDQGVKKIHPSEKIMRP 108
QY 134 QITLFGNDEG--GLMATQVTLSDHGFRMMPVAVYATACELNMLNMPHDKSKCKLIGS 191
DB 109 DLVLYNNADDFALVKTKVLLDYTGHIWTTPAIFKSYCEIIVTHRPFDQNCSSMKLGT 168
QY 192 W---GLKVVLPENGSTRAGSLDHDLDVQSPSEWELVDSRA--HFVSQDYGG-----YME 239
DB 169 WYDGSVVAVINPES---DQPDLSNPMESGEWIKESKGMKHW--FYACCPSTPIYLD 220
QY 240 --YTLTAQRSSMTAVIYTPASCIIVIALSAFMPHPHMGKIKIMINGLIIVIAFLMAY 297
DB 221 ITHFVWQRLPLVFIYVILIPCLLFSFLGLVLYLPD--SGEKMTLSISLVLYLVEFL 279
QY 298 FAQLLPVLSNNTPLVAVFYSTSLIYLSVSTIVEVLVLYLANGKHKR-----LPEALRL 352

```

```

Db 280 IVELIPSTSSAVPLIGKMYLFTMVFIASIIITVIYI---NTHRSRSTHVPKWKV 335
QY 353 LMGHIGTWLLSVSTGTGSAQKTKEMDEHPY-EEADQESS-----PLG-----INH 400
Db 336 F---IDTIPNIMFSTMKRPSREK---ODKKTFTEDIDISDTSKRGPPMKGFHSPLKH 389
QY 401 TEVPGA-----KANQ-----FDMLLATAVDRISFVSLAFLI 434
Db 390 PEVKAIEGKIYIAETMKSDESNNAAEEMKYVAMVMDHILLAVFNLVCIIT 440

RESULT 3
ACH2_RAT 3
ID ACH2_RAT STANDARD: PRT: 511 AA.
AC P12389; 008952;
DT 01-OCT-1989 (Rel. 12, Created)
DT 15-JUL-1998 (Rel. 36, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Neuronal acetylcholine receptor protein, alpha-2 chain precursor.
GN CHRNA2 OR ACRA2.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sclurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
RP SEQUENCE FROM N.A.
RX STRAIN=SPRAGUE-DAWLEY; TISSUE=Brain;
RX MEDLINE=88178113; PUBMED=2832952;
RA Wada K., Ballivet M., Boulter J., Connolly J.G., Wada E.,
RA Denaris E.S., Swanson L.W., Heinemann S.F., Patrick J.;
RT "Functional expression of a new pharmacological subtype of brain
RT nicotinic acetylcholine receptor.";
RL Science 240:330-334(1988).
[2]
RN REVISIONS.
RP STRAIN=SPRAGUE-DAWLEY; TISSUE=Brain;
RC Boulter J.;
RL Submitted (JUN-1997) to the EMBL/Genbank/DBJ databases.
CC -1- FUNCTION: AFTER BINDING ACETYLCHOLINE, THE ACHR RESPONDS BY AN
CC EXTENSIVE CHANGE IN CONFORMATION THAT AFFECTS ALL SUBUNITS AND
CC LEADS TO OPENING OF AN ION-CONDUCTING CHANNEL ACROSS THE PLASMA
CC MEMBRANE.
CC -1- SUBUNIT: NEURONAL ACHR SEEMS TO BE COMPOSED OF TWO DIFFERENT TYPES
CC OF SUBUNITS: ALPHA AND NON-ALPHA (BETA). ALPHA-2 SUBUNIT CAN BE
CC COMBINED TO BETA-2 OR BETA-4 TO GIVE RISE TO FUNCTIONAL RECEPTORS.
CC -1- SUBCELLULAR LOCATION: Integral membrane protein.
CC -1- SIMILARITY: BELONGS TO THE LIGAND-GATED IONIC CHANNELS FAMILY.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@sib-sib.ch).
CC -----
DR EMBL: L10077; AAB60900.1; -
DR EMBL: M20297; AAA40664.1; -
DR EMBL: M20292; AAA40664.1; JOINED.
DR EMBL: M20293; AAA40664.1; JOINED.
DR EMBL: M20294; AAA40664.1; JOINED.
DR EMBL: M20295; AAA40664.1; JOINED.
DR EMBL: M20296; AAA40664.1; JOINED.
DR InterPro: IPR00188; GABA_A_receptor.
DR InterPro: IPR001175; Neur_chan.
DR Pfam: PF02931; Neur_chan_LBD; 1.
DR Pfam: PF02932; Neur_chan_memb; 1.
DR PRINTS: PR00252; NRIONCHANNEL.
DR PROSITE: PS00336; NEURORTRION_CHANNEL; 1.
KW Postsynaptic membrane; Ionic channel; Glycoprotein; Signal;
KW Transmembrane; Multigene family.
FT SIGNAL 1 27
FT CHAIN 28 511 NEURONAL ACETYLCHOLINE RECEPTOR PROTEIN,

```

```

FT DOMAIN 28 241 ALPHA-2 CHAIN.
FT TRANSMEM 242 266 EXTRACELLULAR.
FT TRANSMEM 274 292 POTENTIAL.
FT TRANSMEM 308 329 POTENTIAL.
FT DOMAIN 330 484 CYTOPLASMIC.
FT TRANSMEM 485 503 POTENTIAL.
FT DOMAIN 391 402 POLY-GIL.
FT DISULFID 160 174 BY SIMILARITY.
FT DISULFID 224 225 ASSOCIATED WITH RECEPTOR ACTIVATION
FT CARBOHYD 56 56 (BY SIMILARITY).
FT CARBOHYD 106 106 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 212 212 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CONFLICT 494 494 N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ SEQUENCE 511 AA; 58611 MW; 3824E83BBD1D613B CRC64;

Query Match 18.3%; Score 423; DB 1; Length 511;
Best Local Similarity 24.5%; Pred. No. 1.6e-26;
Matches 126; Conservative 94; Mismatches 167; Indels 128; Gaps 18;

QY 20 LQMLGMLMGLTSPGATATADPKNANYKALDRHAGLFTYDSDVOPFOGTPTNVSL 79
Db 8 LQFWTHLYLWCLLVLP---AVLQOGSHTHARDRLFKLHFGYNNMARPV-----PNTSD 59
QY 80 ENAVVY-----IDELNGKLTHCWLNLNRDEERYWQSQYDNTQITLKSEVWT 132
Db 60 VVIVERGLSTIAOLIDVENQMMNTNVMKQEMNBYKILWMDAEFGNVTSLRVPSEMTWI 119
QY 133 POTTENGEGLMAETQVTLSH---DGHFRMPAPVAVYAGELMNLMPHDKQCKIKI 189
Db 120 PDIYLYNNADGE-FAVTHMTKAHLFEFTGVHWPAPATKSSSIDVTFPPFDQCKMKF 178
QY 190 GSW---GLKVLDPENGTAAGESLDHDDLVQSPMEIVDSRAHVSQDY-----YGMEX 240
Db 179 GSWTYDKAKIDLEQ---MERTVDLKDWESEGMALINATGYNSKKYDCCAEIYPDVY 234
QY 241 TLTQRRSMYAVITYTPASCIIVIALSFMLPRPHGSGKIMINGILLIIVIAFLMYPQ 300
Db 235 YVIRRLPLFYIINLIIPCLLISCLIVLVEYLPSSC-GEKILICSVLSTLVFLLITE 293
QY 301 LPLVSNTPLVVIFSTSLYLSYVIEVLYVLTATGKHK-----RLPALRKLHG 355
Db 294 ITPSTSLVPLIGETILFMTIVTSLVTVLVN---HRSSTHMPMVRVALLG 349
QY 356 HLGTVLLS-----VESTGESQAERTKEMDEHPYEADQES 393
Db 350 RVPRLMNMRLPRPMEHLGSPDLKSPSYHMLETNMDAGERETEE-----EEEDEN 403
QY 394 -----SPGINH-----TEVPGAANOF----- 411
Db 404 ICVCAGLPDSSMGVLYGHGGLHRLAMEPRTKTP-SQASEILLSPQIOKALEGVHIADRL 462
QY 412 -----DMALATAVDRISFVSLAFLI 435
Db 463 RSEDADSVKEDMKYVAMVVDRI---FLMLFTIV 493

RESULT 4
ACH4_HUMAN 4
ID ACH4_HUMAN STANDARD: PRT: 627 AA.
AC P43681;
DT 01-NOV-1995 (Rel. 32, Created)
DT 01-FEB-1996 (Rel. 33, Last sequence update)
DT 01-MAR-2002 (Rel. 41, Last annotation update)
DE Neuronal acetylcholine receptor protein, alpha-4 chain precursor.
GN CHRNA4 OR ACRA4.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
OX NCBI_TaxID=9606;
RN [1]

```

RP SEQUENCE FROM N.A.
 RC TISSUE=Brain;
 RX MEDLINE=95237608; PubMed=7721089;
 RA Monteggia L.M., Gopalakrishnan M., Touna E., Idler K.B., Nash N.,
 RA Aretic S.P., Sullivan J.P., Giordano T.;
 RT "Cloning and transient expression of genes encoding the human alpha-4
 RT and beta-2 neuronal nicotinic acetylcholine receptor (nAChR)
 RT subunits.";
 RL Gene 155:189-193(1995).
 RN [2]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=96430009; PubMed=8833159;
 RA Steinlein O.K., Welland S., Stood J., Propping P.;
 RT "Exon-intron structure of the human neuronal nicotinic acetylcholine
 RT receptor alpha 4 subunit (CHRNA4).";
 RL Genomics 32:289-294(1996).
 RN [3]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Hippocampus;
 RX MEDLINE=97062879; PubMed=8906617;
 RA Elliott K.J., Ellis S.B., Berckhan K.J., Urrutia A.,
 RA Chavez-Noriega L.E., Johnson E.C., Velicelcebi G., Harpold M.M.;
 RT "Comparative structure of human neuronal alpha 2-alpha 7 and beta
 RT 2-beta 4 nicotinic acetylcholine receptor subunits and functional
 RT expression of the alpha 2, alpha 3, alpha 4, alpha 7, beta 2, and
 RT beta 4 subunits.";
 RL J. Mol. Neurosci. 7:217-228(1996).
 RN [4]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=97162233; PubMed=9009220;
 RA Groto Kormelink P.J., Luyten W.H.M.L.;
 RT "Cloning and sequence of full-length cDNAs encoding the human neuronal
 RT nicotinic acetylcholine receptor (nAChR) subunits beta3 and beta4 and
 RT expression of seven nAChR subunits in the human neuroblastoma cell
 RT line SH-SY5Y and/or IMR-32.";
 RL FEBS Lett. 400:309-314(1997).
 RN [5]
 RP SEQUENCE FROM N.A.
 RA Deloukas P., Matthews L.H., Ashurst J., Burton J., Gilbert J.G.R.,
 RA Jones M., Stavrides G., Almeida J.P., Babbage A.K., Baguley C.L.,
 RA Bailey J., Barlow K.F., Bates K.N., Beard L.M., Beare D.M.,
 RA Beasley O.P., Bird C.P., Blakey S.E., Bridgeman A.M., Brown A.J.,
 RA Buck D., Buttill W., Butler A.P., Carder C., Carter N.P., Clee C.M.,
 RA Chapman J.C., Clamp M., Clark G., Clark L.N., Clark S.Y., Clegg C.M.,
 RA Clegg S., Cobley V.E., Collier R.E., Connor R., Corby N.R.,
 RA Coulson A., Coville G.J., Deedman R., Dhami P., Dunn M.,
 RA Ellington A.G., Frankland J.A., Fraser A., French L., Garner P.,
 RA Grafham D.V., Griffiths C., Griffiths M.N.D., Gwilliam R., Hall R.E.,
 RA Hammond S., Harley J.L., Heath P.D., Ho S., Holden J.L., Howden P.J.,
 RA Huckle E., Hunt A.R., Hunt S.E., Jekosch K., Johnson C.M., Johnson D.,
 RA Kay M.P., Kimberley A.M., King A., Knights A., Laird G.K., Lawlor S.,
 RA Leharshlaho M.H., Leverisha M., Lloyd C., Lloyd D.M., Lovell J.D.,
 RA Marsh V.L., Martin S.L., McConachie L.J., McIlroy K., McMurray A.A.,
 RA Oliver S., Mistry D., Moore M.J.F., Mullikin J.C., Nickerson T.,
 RA Oliver K., Parker A., Patel R., Pearce T.A.V., Peck A.I.,
 RA Phillimore B.J.C.T., Prathalingam S.R., Plumb R.W., Ramsay H.,
 RA Ruce C.M., Ross M.T., Scott C.E., Sehra H.K., Showkhen R., Sims S.,
 RA Stuce C.D., Smith M.L., Soderlund C., Steward C.A., Sulston J.E.,
 RA Swann M., Symamore N., Taylor R., Tee L., Thomas D.W., Thorpe A.,
 RA Tracey A., Tromans A.C., Vaudin M., Wall M., Wallis J.M.,
 RA Whitehead S.L., Whitaker P., Willey D.L., Williams L., Williams S.A.,
 RA Wilming L., Wray P.W., Hubbard T., Durbin R.M., Bentley D.R., Beck S.,
 RA Rogers J.;
 RT "The DNA sequence and comparative analysis of human chromosome 20.";
 RL Nature 414:865-871(2001).
 RN [6]
 RP SEQUENCE OF 26-627 FROM N.A.
 RC TISSUE=Brain;
 RA Mamalaki A., Remoundos M., Tzartos S.;
 RL Submitted (May-1995) to the EMBL/GenBank/DBJ databases.
 RN [7]
 RP VARIANT ENFL1 PHE-280
 RX MEDLINE=96024658; PubMed=7550350;

RA Steinlein O.K., Mulley J.C., Propping P., Wallace R.H., Phillips H.A.,
 RA Sutherland G.R., Scheffer I.E., Berkovic S.F.;
 RT "A missense mutation in the neuronal nicotinic acetylcholine receptor
 RT alpha-4 subunit is associated with autosomal dominant nocturnal
 RT frontal lobe epilepsy.";
 RL Nat. Genet. 11:201-203(1995).
 CC -1- FUNCTION: AFTER BINDING ACETYLCHOLINE, THE AChR RESPONDS BY AN
 CC EXTENSIVE CHANGE IN CONFORMATION THAT AFFECTS ALL SUBUNITS AND
 CC LEADS TO OPENING OF AN ION-CONDUCTING CHANNEL ACROSS THE PLASMA
 CC MEMBRANE.
 CC -1- SUBUNIT: NEURONAL AChR SEEMS TO BE COMPOSED OF TWO DIFFERENT
 CC TYPE OF SUBUNITS: ALPHA AND NON-ALPHA (BETA).
 CC -1- SUBCELLULAR LOCATION: Integral membrane protein.
 CC -1- DISEASE: Defects in CHRNA4 are the cause of autosomal dominant
 CC nocturnal frontal lobe epilepsy type 1 (ENFL1 or ADNFLE). It is a
 CC disease characterized by clusters of motor seizures during sleep.
 CC -1- SIMILARITY: BELONGS TO THE LIGAND-GATED IONIC CHANNELS FAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL: L35901; AAA64743.1; -;
 DR EMBL: X89741; CAA61893.1; -;
 DR EMBL: X89742; CAA61893.1; JOINED.
 DR EMBL: X89743; CAA61893.1; JOINED.
 DR EMBL: X89744; CAA61893.1; JOINED.
 DR EMBL: X89745; CAA61893.1; JOINED.
 DR EMBL: X89746; CAA61893.1; JOINED.
 DR EMBL: U62433; AAB40111.1; -;
 DR EMBL: X08421; CAA69698.1; -;
 DR EMBL: AL121827; CAC36119.1; -;
 DR EMBL: X87629; CAA60959.1; -;
 DR MIM: 118504; -;
 DR MIM: 600513; -;
 DR InterPro: IPR000188; GABAA_receptor.
 DR InterPro: IPR001175; Neur_channel.
 DR Pfam: PF02931; Neur_chan_LBD; 1.
 DR Pfam: PF02932; Neur_chan_memb; 1.
 DR PRINTS: PR00252; NRIONCHANNEL.
 DR PROSITE: PS00236; NEUROTR_IOM_CHANNEL; 1.
 KW Postsynaptic membrane; Ionic channel; Glycoprotein; Signal;
 KW Transmembrane; Multigene family; Disease mutation.
 FT SIGNAL 1
 FT CHAIN 28
 FT FT 29 627
 FT DOMAIN 29 242
 FT TRANSMEM 243 267
 FT TRANSMEM 275 293
 FT TRANSMEM 309 330
 FT DOMAIN 331 600
 FT TRANSMEM 601 619
 FT DISULFID 161 175
 FT DISULFID 225 226
 FT FT 280 280
 FT VARIANT 280 280
 FT FT
 FT FT
 SQ SEQUENCE 627 AA; 69957 MW; B3AOC0151E5A2A8 CRC64;

Query Match 18.1%; Score 418; DB 1; Length 627;
 Best Local Similarity 29.8%; Pred. No. 5.2e-26;
 Matches 102; Conservative 67; Mismatches 139; Indels 34; Gaps 9;
 QY 11 VSGGLPLQLQMLGMLMLGTSVPGATATADPNANAVKALDRHAGLFTNYSDVQPVF 70
 Db 3 LGFGGAPRLRLPLLLLLGTGLRRSSHYET-----RAHAEERLKKLFEGYKNWSPV- 55
 QY 71 QGTFPNSLEAVYTY-----IDIDELNGKLTTHCWLNRWDERVWQPSQYDNTIQT 123

```

Db 56 ----ANISDVVLVRGLSLAQILIDVDEKQNMNTVNWVQKQEMHDKYLRMDPADYEIVTSI 111
QY 124 TLKSESVWTPQTTLFNGDEGGGLMAE--IQVYTLSHDGHFWMPRAYTATVCELNMLNWPMD 181
Db 112 RIPELIMPRDIDLVLNMAQDPAVYHLLTKAHLFHDGRVQMTPEIAYKSSCIDYVEFFPPD 171
QY 182 KQSCCKLKISW---GLKVVLPENGTARGESLDHDDLVSQPEMEIVDSRAHFVSODY---- 234
Db 172 QQNCITKRESSTWYDKAKIDL-VNMSR---VDQLDFWEGEYVYDAVCTYITRKYECCA 227
QY 235 --YGYMEYTLTAORSSMYTAVIYTPASCIVILALSAFWLPPHMGGEKIMINGLLIIVIA 292
Db 228 ELYPDITVFAVIRLPLFYTIMLIPCLLISCLTVLVEFLPSPSC-GEKITLICISVLSTLT 286
QY 293 AFLMFAQLLPVLSNNTPLVVFYSTLSLYSVSTIVEVLV 334
Db 287 VELLITELIPSTSLVPLIGELYLTMTLFTVLSIVITVEVL 328

```

RESULT 5

```

ACH2_HUMAN
ID ACH2_HUMAN STANDARD: PRT: 529 AA.
AC Q15822; Q9H4Q3;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 01-MAR-2002 (Rel. 41, Last annotation update)
DE Neuronal acetylcholine receptor protein, alpha-2 chain precursor.
GN CHRNA2.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
OX NCBI_TaxId=9606;
RX TISSUE=Hypothalamus;
RX MEDLINE=97062879; PubMed=8906617;
RA Elliott K.J., Ellis S.B., Borchan K.J., Urrutia A.,
RA Chavez-Noriega L.E., Johnson E.C., Velicelceti G., Harpold M.M.;
RT "Comparative structure of human neuronal alpha 2, alpha 7 and beta
RT 2-beta 4 nicotinic acetylcholine receptor subunits and functional
RT beta 4 subunits.";
RT J. Mol. Neurosci. 7:217-228(1996).
RN [2]
RP SEQUENCE FROM N.A.
RA Groot Kormelink P.J.;
RL Submitted (JAN-1998) to the EMBL/GenBank/DBJ databases.
RN [3]
RP SEQUENCE FROM N.A.
RA Blechschmidt K., Rosenthal A.;
RL Submitted (OCT-2000) to the EMBL/GenBank/DBJ databases.
CC -1- FUNCTION: AFTER BINDING ACETYLCHOLINE, THE ACHR RESPONDS BY AN
CC EXTENSIVE CHANGE IN CONFORMATION THAT AFFECTS ALL SUBUNITS AND
CC LEADS TO OPENING OF AN ION-CONDUCTING CHANNEL ACROSS THE PLASMA
CC MEMBRANE.
CC -1- SUBUNIT: NEURONAL ACHR SEEMS TO BE COMPOSED OF TWO DIFFERENT TYPES
CC OF SUBUNITS: ALPHA AND NON-ALPHA (BETA). ALPHA-2 SUBUNIT CAN BE
CC COMBINED TO BETA-2 OR BETA-4 TO GIVE RISE TO FUNCTIONAL RECEPTORS.
CC -1- SUBCELLULAR LOCATION: Integral membrane protein.
CC -1- SIMILARITY: BELONGS TO THE LIGAND-GATED IONIC CHANNELS FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation-
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@sib-sib.ch).
CC -----
DR EMBL; U62431; AAB40109.1; -
DR EMBL; Y16281; CAA76154.1; -
DR EMBL; AF311103; -; NOT_ANNOTATED_CDS.

```

```

DR MIM; 118502; -
DR InterPro: IPR000188; GABA_receptor.
DR InterPro: IPR001175; Neur_channel.
DR Pfam; PF02931; Neur_chan_LBD; 1.
DR Pfam; PF02932; Neur_chan_membr; 1.
DR PRINTS; PR00252; NRIONCHANNEL.
DR PROSITE; PS00236; NEURON_ION_CHANNEL; 1.
DR Postsynaptic membrane; Ionic channel; Glycoprotein; signal;
DR Transmembrane; Multigene family.
FT SIGNAL 1
FT CHAIN 26
FT 27 529
FT DOMAIN 27 264
FT TRANSMEM 265 289
FT TRANSMEM 297 315
FT TRANSMEM 331 352
FT DOMAIN 353 502
FT TRANSMEM 503 521
FT TRANSMEM 183 197
FT DISULFID 247 248
FT CARBOHYD 79 79
FT CARBOHYD 129 129
FT CARBOHYD 235 235
FT CONFLICT 125 125
FT SEQUENCE 529 AA; 59735 MW; 7F512B06CDD9A9FD CRC64;

```

Query Match 18.0%; Score 416; DB 1; Length 529;
 Best Local Similarity 27.6%; Pred. No. 6e-26;
 Matches 109; Conservative 75; Mismatches 153; Indels 58; Gaps 12;

```

QY 13 GPCPLLLQML-MGMLMLQTLSPGATATADKNA-----NKAIDR 53
Db 2 GSPCVFLSFTKLSLMMWLLTLPAGEGEAKRPPRAGDPLSSPPTALPQGGSHETEDR 61
QY 54 LHAQLETNVSDVOPFQGTPTNVSLSEVYVY-----IDDELNGKLTTCWNLNMR 106
Db 62 LKHLHFGYINRRARVY-----PNTSDVYIVRGSLSLAQILIDVDEKQNMNTVNWVQKQEMS 116
QY 107 DEERWQPSQYDNTIQTLKSESVWTPQTTLFNGDEGGGLMAETQYTLN--DGHFRMP 163
Db 117 DYKLRWNPADFCNITSLRVSEMIWIPDIVLNNDGE-FAYTHMTKAHLFSTGVHMYV 175
QY 164 PAVYTAFCCLNMLNMPHDKQSCCKLKISW---GLKVVLPENGTARGESLDHDDLVSQPEW 220
Db 176 PAIYKSSCSIDYVTFPPDOONCKMKFGSWTYDKAKIDLEQ---MEQIVDLKDYESGEM 231
QY 221 EIVDSRAHFVSODY-----YGYMEYTLTAORSSMYTAVIYTPASCIVILALSAFWLPP 274
Db 232 AIVNATGYNYSKKYDCCATITPDYTAFAVIRLPLFYTIMLIPCLLISCLTVLVEFLPS 291
QY 275 HMGGEKIMINGLLIIVIAFLMFAQLLPVLSNNTPLVVFYSTLSLYSVSTIVEVLV 334
Db 292 DC-GEKITLICISVLSTVLELLITELIPSTSLVPLIGELYLTMTLFTVLSIVITVEVL 350
QY 335 YLATGKHKR-----RLPEALRKLLHCHLCTWLLLS 364
Db 351 NV-----HHRSPSTHMPHWYRGALLGVPRWLLMN 381

```

RESULT 6

```

ACHA_RAT
ID ACHA_RAT STANDARD: PRT: 457 AA.
AC P25108;
DT 01-MAY-1992 (Rel. 22, Created)
DT 01-MAY-1992 (Rel. 22, Last sequence update)
DT 01-OCT-1996 (Rel. 34, Last annotation update)
DE Acetylcholine receptor protein, alpha chain precursor.
GN ACHRA.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.

```

OX NCBI_TaxID=10116;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Muscle;
 RX MEDLINE=91099317; PubMed=1702709;
 RA Witzmann V., Steh E., Barg B., Konno T., Koenen M., Kues W.,
 RA Chlad M., Hofmann M., Sakmann B.;
 RT "Primary structure and functional expression of the alpha-, beta-,
 RT gamma-, delta- and epsilon-subunits of the acetylcholine receptor
 RT from rat muscle.";
 RL Eur. J. Biochem. 194;437-448(1990).
 CC -1- FUNCTION: AFTER BINDING ACETYLCHOLINE, THE ACHR RESPONDS BY AN
 CC EXTENSIVE CHANGE IN CONFORMATION THAT AFFECTS ALL SUBUNITS AND
 CC LEADS TO OPENING OF AN ION-CONDUCTING CHANNEL ACROSS THE PLASMA
 CC MEMBRANE.
 CC -1- SUBUNIT: PENTAMER OF TWO ALPHA CHAINS, AND ONE EACH OF THE BETA,
 CC DELTA, AND GAMMA (IN IMMATURE MUSCLE) OR EPSILON (IN MATURE
 CC MUSCLE) CHAINS.
 CC -1- SUBCELLULAR LOCATION: Integral membrane protein.
 CC -1- SIMILARITY: BELONGS TO THE LIGAND-GATED IONIC CHANNELS FAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL: X74832; CA52826.1; -
 DR PIR: S13872; S13872.
 DR InterPro: IPR000188; GABAA_receptor.
 DR Pfam: PF02931; Neur_chan_LBD; 1.
 DR Pfam: PF02932; Neur_chan_memb; 1.
 DR PRINTS: PR00252; NRIONCHANNEL.
 DR PROSITE: PS00236; NEUROTR_IOM_CHANNEL; 1.
 DR Receptor; Postsynaptic membrane; Ionic channel; Glycoprotein; Signal;
 KW Transmembrane.
 KM
 FT SIGNAL 1 20
 FT CHAIN 21 457
 FT DOMAIN 21 230
 FT TRANSMEM 231 235
 FT TRANSMEM 263 281
 FT TRANSMEM 297 316
 FT DOMAIN 317 428
 FT TRANSMEM 429 447
 FT DISULFID 148 162
 FT DISULFID 212 213
 FT CARBOHYD 161 161
 FT SEQUENCE 457 AA; 51866 MW; 776AE3B8DF8F68B3 CRC64;
 SO
 Query Match 17.9%; Score 414.5; DB 1; Length 457;
 Best Local Similarity 25.4%; Pred. No. 6.6e-26;
 Matches 118; Conservative 99; Mismatches 174; Indels 73; Gaps 19;
 OY 20 LOMLGMGLGITSVPGATADPKNANVKALDLHGLFNNYSDQPV---FQGIPTN 76
 Db 1 MELTAVILLGLGCSA-STVLASEHET-----RLVAKLEKDYSSVVRPGDREHREIVQT 52
 OY 77 VLEAVVTVYIDIDELNGKLTTHCWLNRWDEERWOPSOVDNTQTLLKSEVWTPQIT 136
 Db 53 VGLQL-IQLINVDENVNITTVNRLKQGWYDNLKAMPDYGKAKHIBEKRTREDVY 111
 OY 137 LFNQDEG--GLAETQVTLSDHGFRMPAPVATYACELANLNPHDKQSCIKLIGSW-- 192
 Db 112 LYNNADEDFAIKFTKLLDYTGHTWTPTPAIFKSYCEIIVTHPPDEQNSMKLGTWY 171
 OY 193 --GLKAVLPENGTAEGSLDHDLDVQSPWEIVDSRA--HFVSDYYG-----YME--Y 240

Db 172 DGSVAINPES-----DQPLDSNFMESGENVKEARQWKHNV---FYSCCPNTPYDITY 223
 OY 241 TLTPARRSSMTAVIYTPASCTVILALSAFLPFPNMGKIMINGLLIYIAAFMTFAQ 300
 Db 224 HFVQRLPLVFIYVILIPCLLFSFLTSLVFLYLPD-SGEMNTLSISLSTVFLVIVE 282
 OY 301 LLLPVSNTPLVVIFSTSLYSVSTIVEVLYLYATGKHRR-----LPEALRKLHG 355
 Db 283 LIPSTSSAVPLICKYMLFTWVFYIASIITIVIVI---NTHHRSPTTHMPEWVRKVFID 338
 OY 356 HLGFWLLSVFSTGCEQAK---TKEMDHPVEADEQESSPLG-----INHEVGA- 406
 Db 339 TTPIMFSTMKRPSRDKQERITFEDIDS--DISKPGPPPGFHSPLIKPEVSAI 396
 OY 407 -----KAND-----FDNALLATAVDRISFVSFLAFL 434
 Db 397 EGVKVIETMKSDQESNNASEWKYVAMVDHLLGVFMLVCL 440
 RESULT 7
 ACH6_RAT STANDARD; PRT; 493 AA.
 AC PA3143;
 DT 01-NOV-1995 (Rel. 32, Created)
 DT 01-NOV-1995 (Rel. 32, Last sequence update)
 DT 15-JUL-1999 (Rel. 38, Last annotation update)
 DE Neuronal acetylcholine receptor protein, alpha-6 chain precursor.
 GN CHRN6 OR ACRA6.
 OS Rattus norvegicus (Rat).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
 OX NCBI_TaxID=10116;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Boulter J.;
 RU Submitted (XXX-1992) to the EMBL/GenBank/DBJ databases.
 CC -1- FUNCTION: AFTER BINDING ACETYLCHOLINE, THE ACHR RESPONDS BY AN
 CC EXTENSIVE CHANGE IN CONFORMATION THAT AFFECTS ALL SUBUNITS AND
 CC LEADS TO OPENING OF AN ION-CONDUCTING CHANNEL ACROSS THE PLASMA
 CC MEMBRANE.
 CC -1- SUBCELLULAR LOCATION: Integral membrane protein.
 CC -1- SIMILARITY: BELONGS TO THE LIGAND-GATED IONIC CHANNELS FAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL: L08227; AAA41674.1; -
 DR HSSP: P01270; 1HPH.
 DR InterPro: IPR000188; GABAA_receptor.
 DR InterPro: IPR001175; Neur_channel.
 DR Pfam: PF02931; Neur_chan_LBD; 1.
 DR Pfam: PF02932; Neur_chan_memb; 1.
 DR PRINTS: PR00252; NRIONCHANNEL.
 DR PROSITE: PS00236; NEUROTR_IOM_CHANNEL; 1.
 DR Receptor; Postsynaptic membrane; Ionic channel; Glycoprotein; Signal;
 KW Transmembrane; Multigene family.
 KM
 FT SIGNAL 1 30
 FT CHAIN 31 493
 FT DOMAIN 31 240
 FT TRANSMEM 241 265
 FT TRANSMEM 272 290
 FT TRANSMEM 306 327
 FT DOMAIN 328 464
 FT TRANSMEM 465 484
 FT DISULFID 158 172
 FT DISULFID 222 223
 FT
 By Similarity.
 NEURONAL ACETYLCHOLINE RECEPTOR PROTEIN,
 ALPHA-6 CHAIN.
 EXTRACELLULAR (POTENTIAL).
 POTENTIAL.
 POTENTIAL.
 CYTOPLASMIC (POTENTIAL).
 POTENTIAL.
 BY SIMILARITY.
 ASSOCIATED WITH RECEPTOR ACTIVATION
 (BY SIMILARITY).

Db 289 SAVFLIKYMLFTWVFIASIIITVIYI---NTHHSPSTHINPEWVKVEIDTIPNIM 344

QY 362 LLSVSTTGSQAEK---TKEMDEHPYEADDEQESSPLG-----INHTVEPGA----- 406

Db 345 FFSFMKRPNRKQKRIFTEDIDIS--DISGKPGPPPMGFSPILKHPVKSALGVKYI 402

QY 407 ----KANO-----FDWALLATAVDRISFVFSFSLATLI 434

Db 403 AETWKSDESNNAEWEKVVYAMVDHILLGVFMLVCLI 440

RESULT 9

ACHO_CHICK ID ACHO_CHICK STANDARD; PRT; 455 AA.

AC P43679;

DT 01-NOV-1995 (Rel. 32, Created)

DT 16-OCT-2001 (Rel. 40, last sequence update)

DT 16-OCT-2001 (Rel. 40, last annotation update)

DE Neuronal acetylcholine receptor protein, beta-3 chain precursor.

OS Gallus gallus (Chicken).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae; Gallus.

OC NCBI_TaxID=9031;

OX [1]

RN SEQUENCE FROM N.A.

RP STRAIN-WHITE LECHORN;

RX MEDLINE=95155414; PubMed=7852408;

RA Hernandez M.C., Erkman L., Matter-Sadzinski L., Roztocil T., Ballivet M., Matter J.M., "Characterization of the nicotinic acetylcholine receptor beta 3 gene. Its regulation within the avian nervous system is effected by a promoter 143 base pairs in length.";

RT J. Biol. Chem. 270:3224-3233(1995).

RL [2]

RN REVISION TO 331.

RA Ballivet M.;

RL Submitted (OCT-1999) to the EMBL/GenBank/DBJ databases.

CC -1- FUNCTION: AFTER BINDING ACETYLCHOLINE, THE ACHR RESPONDS BY AN EXTENSIVE CHANGE IN CONFORMATION THAT AFFECTS ALL SUBUNITS AND LEADS TO OPENING OF AN ION-CONDUCTING CHANNEL ACROSS THE PLASMA MEMBRANE.

CC -1- SUBUNIT: NEURONAL ACHR SEEMS TO BE COMPOSED OF TWO DIFFERENT TYPE OF SUBUNITS: ALPHA AND BETA.

CC -1- SUBCELLULAR LOCATION: Integral membrane protein.

CC -1- TISSUE SPECIFICITY: RELATIVELY ABUNDANT IN THE DEVELOPING RETINA AND IN THE TRIGEMINAL GANGLION.

CC -1- SIMILARITY: BELONGS TO THE LIGAND-GATED IONIC CHANNELS FAMILY.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).

CC -----

CC EMBL; X83739; CAB59814.1; -

DR InterPro: IPR000188; GABAA_receptor.

DR InterPro: IPR000175; Neur_channel.

DR Pfam: PF02931; Neur_chan_Lbd; 1.

DR Pfam: PF02932; Neur_chan_memb; 1.

DR PRINTS: PR00252; NRIONCHANNEL.

DR PROSITE: PS00236; NEUROTR_ION_CHANNEL; 1.

KW postsynaptic membrane; Ionic channel; Glycoprotein; Signal;

KW Transmembrane; Multigene family

FT SIGNAL 1 20

FT CHAIN 1 455

FT DOMAIN 21 229

FT TRANSMEM 230 254

FT TRANSMEM 262 279

FT TRANSMEM 296 317

FT DOMAIN 318 425

FT TRANSMEM 426 444

FT DISULFID 150 164

FT CARBOHYD 48 48

FT CARBOHYD 163 163

SO SEQUENCE 455 AA; 52276 MW; 3D7A7ARA77D8511A CRC64;

Query Match 17.7%; Score 410; DB 1; Length 455;

Best local Similarity 26.3%; Pred. No. 1.5e-25;

Matches 115; Conservative 85; Mismatches 167; Indels 70; Gaps 17;

QY 52 DRCHAGLFTWVDSQVDFVFGTPT-NVSLDMVYT-YIDIDELNGKLTTHCWLIRWDEE 109

Db 27 DALRHLEFQCYQKWKVRVENSNDITIKVLFGLKISQLDVDEKQMLMTNWMLAQEWMDK 86

QY 110 RVMPQSDYDNTQTTLTKSSVWTPQITLF-NGD--EGGIMAEFGVYLSHDGHRMRPPA 165

Db 87 LSMNPEEYGGITATRVSESILMPDIVLFENADGREGSLM--TKATVKNVGVQWMPA 144

QY 166 VYTAVCELMNMPHDKOSCKLIGSWG-----LKVLPENGTARGESLDHDLVQSPW 220

Db 145 SYKSSCTMELFPFPDQNSMKFGSWTYGSDWDLVD-----ENVDTKDFPDNGEM 198

QY 221 EYDSRAHEVSQD---YGYMEYTLTAQRSSMYTAVIYTPASCIYIALSAFWLPPHM 276

Db 199 EILNAKMGKGNRKDGISTYFPFVYSFVLRRLPLEYTLFLIPCLGSLFLVLFVYLPSPD- 257

QY 277 GGEKIMINGLLIYIAFLWYFNQDLVLSNNPLVYIFSTSLYLSVSTIYEVLYL 336

Db 258 EGKLSLSTSVLSLVLLVLEIEIIPSSKVIPLIGEVLYLFIETVLSIYVFIYV 317

QY 337 ATGKHKR-----RLPEALRKLLHGLTWLLSVSTTGSQAE-KT 377

Db 318 ----HHRSSATYHPMAWPKRFLQKLPRL--CMKHVDVRY----SFSDTEKETTLKS 367

QY 378 KEMDEHPYEADDEQESSPLGINHTVEPG-----AKANQF-----DWALLATAVDRIS 424

Db 368 LKPKCKHQAQKDEKVFIAFLERKADSIYRSHVKKDAFIQOVQDMKFAVQVDRI- 426

QY 425 FVSPSLAFLATIRCSV 441

Db 427 ---FLWLFVAVSYGVSV 440

RESULT 10

ACHO_HUMAN ID ACHO_HUMAN STANDARD; PRT; 450 AA.

AC Q9GZ26;

DT 16-OCT-2001 (Rel. 40, Created)

DT 16-OCT-2001 (Rel. 40, last sequence update)

DT 01-MAR-2002 (Rel. 41, last annotation update)

DE Neuronal acetylcholine receptor protein, alpha-10 chain precursor

DE (nicotinic acetylcholine receptor subunit alpha 10) (NACHR alpha 10).

GN CHRNA10 OR NACHRA10.

OS Homo sapiens (Human).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.

OC NCBI_TaxID=9606;

OX [1]

RN SEQUENCE FROM N.A.

RP TISSUE=Skeletal muscle;

RA Sgar F., Charpanlier E., Bertrand S., Walker N., Agnel M., Graham D., Caput D., Bertrand D., Besnard F.;

RT "Cloning and expression of a novel nicotinic alpha10 subunit from human that confers functionality to the alpha9 subunit.";

RL Submitted (MAY-2000) to the EMBL/GenBank/DBJ databases.

RN [2]

RP SEQUENCE FROM N.A.

RC TISSUE=Inner ear, and Tonsil;

RX MEDLINE=21248672; PubMed=11350119;

RA Lustig L.R., Peng H., Hiel H., Yamamoto T., Fuchs P.A.;

RT "Molecular cloning and mapping of the human nicotinic acetylcholine


```

RT receptor alpha10 (CHRNA10).";
RL Genomics 73:272-283(2001).
RN [3]
RP SEQUENCE FROM N.A.
RA Vandenberk I., van der Helm L., Nieuwstraten D.M., Verfallle C.,
RA Kremer A., van der Spek P., Masure S., Hoefnagel E., Yon J.,
RA Groot Kormelink P.J., Luyten W.H., Grantham C.J.;
RT "Identification and functional characterisation of a novel human
RT neuronal nicotinic acetylcholine receptor subunit alpha 10.",
RT submitted (NOV-2000) to the EMBL/Genbank/DBPST databases.
CC -1- FUNCTION: AFTER BINDING ACETYLCHOLINE, THE ACHR RESPONDS BY AN
CC EXTENSIVE CHANGE IN CONFORMATION THAT AFFECTS ALL SUBUNITS AND
CC LEADS TO OPENING OF AN ION-CONDUCTING CHANNEL ACROSS THE PLASMA
CC MEMBRANE.
CC -1- SUBCELLULAR LOCATION: Integral membrane protein.
CC -1- TISSUE SPECIFICITY: Expressed in inner-ear tissue, tonsil,
CC immortalized B-cells, cultured T-cells and peripheral blood
CC lymphocytes.
CC -1- SIMILARITY: BELONGS TO THE LIGAND-GATED IONIC CHANNELS FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@sib-sib.ch).
CC -----
DR EMBL: AJ278118; CAC20435.1; -
DR EMBL: AF199235; AAG00795.2; -
DR EMBL: AF327367; AAK14333.1; -
DR EMBL: AJ295237; CAC16144.1; -
DR MIM: 606372; -
DR InterPro: IPR001188; GABA_A_receptor.
DR InterPro: IPR001175; Neur_channel.
DR Pfam: PF02931; Neur_chan_LBD; 1.
DR Pfam: PF02932; Neur_chan_memb; 1.
DR PROSITE: PS00252; NRIONCHANNEL.
DR PROSITE: PS00236; NEUROTR_ION_CHANNEL; 1.
KW Postsynaptic membrane; Ionic channel; Glycoprotein; Signal;
KW Transmembrane; Multigene family.
FT SIGNAL 1 24
FT CHAIN 25 450
FT FT NEURONAL ACETYLCHOLINE RECEPTOR PROTEIN,
FT ALPHA-10 CHAIN.
FT DOMAIN 25 237
FT TRANSMEM 228 258
FT TRANSMEM 268 288
FT TRANSMEM 302 322
FT DOMAIN 323 428
FT TRANSMEM 429 449
FT TRANSMEM 154 168
FT DISULFID 218 219
FT FT ASSOCIATED WITH RECEPTOR ACTIVATION
FT CARBOHYD 40 40 (BY SIMILARITY).
FT CARBOHYD 56 56 (N-LINKED (GLCNAC. . .) (POTENTIAL).
FT SEQUENCE 450 AA; 49704 MW; 0A1DAD32D9069870 CRC64;
SO

```

```

QY 197 -VLPENGTARGESLDHDDLVQSPMEIVDSRA-----HFVSODYGYMEYTLTAORSS 249
DB 183 DWRP-----RGAASLADVEVENEMRVLCMPARRRULYGCSEPPDYVTLRRRA 237
QY 250 MYTAVIYTPASCVIILALSAFLPPIPMGEKIMINGLLIYIAFLMFAQLLPVLSNNT 309
DB 238 AVVCNLLPCLVILSLAPLAFHLPD-SEGEKSLGVTLALVQLLAESMP-ASEV 295
QY 310 PLVYIFSTSLYSLVSTIVEVLVL-ATGAKRRRLPALKLLGHGLTWLLSVST 368
DB 296 PLIGKYATMTVMVFESTALITLIMNLHYCGSVRPVPMARALLGHRLGLCV---RE 352
QY 369 TGEQAE-KTKEMDEHPYEADDESSPG-----INHTPEPKANQF-- 411
DB 353 REEPGQSGAPPELSPSP-QSEPGAGPPGCPGHEPCLCROELAHV---ATIANFSS 408
QY 412 -----DMALLATAVDRISFVSFLAILAI 437
DB 409 HRAARCHEMDWRKLARVMDRFLAIFFSMALVMSL 443

```

RESULT 11

ID	ACH4_CHICK	STANDARD:	PRT:	622 AA.
AC	P09482;			
DT	01-MAR-1989 (Rel. 10, Created)			
DT	01-MAR-1989 (Rel. 10, Last sequence update)			
DT	16-OCT-2001 (Rel. 40, Last annotation update)			
DE	Neuronal acetylcholine receptor protein, alpha-4 chain precursor.			
OS	Gallus gallus (Chicken).			
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
OC	Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;			
OC	Gallus.			
OX	NCBI_TaxID=9031;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RC	TISSUE=Brain;			
RX	MEDLINE=88283624; PubMed=3267226;			
RA	Nef P., Oneyser C., Alliod G., Couturier S., Ballivet M.;			
RT	"Genes expressed in the brain define three distinct neuronal			
RT	nicotinic acetylcholine receptors.";			
RL	EMBO J. 7:595-601(1988).			
RN	[2]			
RP	MUTAGENESIS OF GLU-289, AND SUBUNITS.			
RX	MEDLINE=91172320; PubMed=2005979;			
RA	Cooper E., Couturier S., Ballivet M.;			
RT	"Pentameric structure and subunit stoichiometry of a neuronal			
RT	nicotinic acetylcholine receptor.";			
RL	Nature 350:235-238(1991).			
CC	-1- FUNCTION: AFTER BINDING ACETYLCHOLINE, THE ACHR RESPONDS BY AN			
CC	EXTENSIVE CHANGE IN CONFORMATION THAT AFFECTS ALL SUBUNITS AND			
CC	LEADS TO OPENING OF AN ION-CONDUCTING CHANNEL ACROSS THE PLASMA			
CC	MEMBRANE.			
CC	-1- SUBUNIT: NEURONAL ACHR SEEMS TO BE COMPOSED OF TWO DIFFERENT			
CC	TYPE OF SUBUNITS: ALPHA AND NON-ALPHA (ALSO CALLED BETA). A			
CC	FUNCTIONAL RECEPTOR SEEMS TO CONSIST OF TWO ALPHA-CHAINS AND			
CC	THREE NON-ALPHA CHAINS.			
CC	-1- SUBCELLULAR LOCATION: Integral membrane protein.			
CC	-1- SIMILARITY: BELONGS TO THE LIGAND-GATED IONIC CHANNELS FAMILY.			
CC	-----			
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration			
CC	between the Swiss Institute of Bioinformatics and the EMBL outstation -			
CC	the European Bioinformatics Institute. There are no restrictions on its			
CC	use by non-profit institutions as long as its content is in no way			
CC	modified and this statement is not removed. Usage by and for commercial			
CC	entities requires a license agreement (See http://www.isb-sib.ch/announce/			
CC	or send an email to license@sib-sib.ch).			
CC	-----			
DR	EMBL: X07348; CAA30285.1; -			
DR	EMBL: X07349; CAA30285.1; JOINED.			
DR	EMBL: X07350; CAA30285.1; JOINED.			
DR	EMBL: X07351; CAA30285.1; JOINED.			

DR EMBL: X07352; CAA30285.1; JOINED.
 DR EMBL: X07359; CAA30285.1; JOINED.
 DR EMBL: AJ250361; CAB59626.1; -.
 DR PIR: S00379; ACCH4N.
 DR InterPro: IPR000188; GABAA_receptor.
 DR InterPro: IPR001175; Neur_channel.
 DR Pfam: PF02931; Neur_chan_LBD; 1.
 DR Pfam: PF02932; Neur_chan_memb; 1.
 DR PRINTS: PR00252; NRIONCHANNEL.
 DR PROSITE: PS00236; NEUROTR-ION_CHANNEL; 1.
 KW Postsynaptic membrane; Ionic channel; Glycoprotein; Signal;
 KW Transmembrane; Multigene family.
 FT SIGNAL 1 23
 FT CHAIN 24 622
 FT DOMAIN 24 262 NEURONAL ACETYLCHOLINE RECEPTOR PROTEIN,
 FT TRANSMEM 238 262 ALPHA-4 CHAIN.
 FT TRANSMEM 270 288 EXTRACELLULAR.
 FT TRANSMEM 304 325
 FT DOMAIN 326 325 CYTOPLASMIC.
 FT TRANSMEM 596 614
 FT DISULFID 156 170
 FT DISULFID 220 221
 FT CARBOHYD 52 52 (BY SIMILARITY).
 FT CARBOHYD 102 102 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT MOTABEN 289 289 E-K; REDUCES CHANNEL CONDUCTANCE BY
 FT HALF.
 SQ SEQUENCE 622 AA; 70837 MW; 290B035893D6485A CRC64;

Query Match 17.6%; Score 407.5; DB 1; Length 622;
 Best Local Similarity 29.4%; Pred. No. 3.6e-25;
 Matches 96; Conservative 73; Mismatches 119; Indels 39; Gaps 10;
 QY 50 ALDLHAGLFTNYSDDVQPGPTNYSLEMYTY-----IDIDELNKLTHCHLN 102
 DB 31 AEEPLKLLKFGYKMSRPV-----ANISDVAVRFGLSIAQLIDVDKKNMTTNVWK 85
 QY 103 LRMDEERWVQPSQYDNTQTLLKSSEVWTPQITLNFDEGLMAEYQVTLSSH--DGHF 159
 DB 86 QEMHDYKLRMPQREYEVNTSIRISBELMRDYLNNADD-FAVHTLTAHLFYDORI 144
 QY 160 RMPFPAVYTAACELNMLNMPHDKOSKIKGSW---GLKVVLPENGARGESLDHDLVQ 216
 DB 145 KMPFPAVYKSSCIDVFEFPDQONCKMKFSGWTYDAKIDL---VSMHSHVDQLDWE 200
 QY 217 SPEPEIYDSRAHEVSQDY-----YGYMEYTLTAQRSSMYTAVIYTPASCIIVILASAF 270
 DB 201 SGEVITINAGVNSKRYEECTETIYPTYSFTIRRLPLEFTYIMLILPCLLSCTLVLF 260
 QY 271 WLPFHMGEKIMINGLLIIVIAFLMFAQLPLVLSNNTPLVVFYSTLLLSVSTIVE 330
 DB 261 YLPSECG-GEKTLTCLISVLSTFLVLLITETIIPSTSLVPLIGRYLFTMLFVLSITIT 319
 QY 331 VLVIYLTATGKRK-----RLPEALRKL 352
 DB 320 VEVLVN-----HHRSPRTHTPMDWVRV 342

RESULT 12
 ACHI_SCHGR STANDARD; PRT; 557 AA.
 AC P23414;
 DT 01-NOV-1991 (Rel. 20, last sequence update)
 DT 01-NOV-1991 (Rel. 20, last sequence update)
 DT 01-FEB-1994 (Rel. 28, last annotation update)
 DE Acetylcholine receptor protein, alpha-L1 chain precursor.
 OS Schistocerca gregaria (Desert locust).
 OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
 OC Phlebotominae; Neoptera; Orthopteroidea; Orthoptera; Caellifera;
 OC Acridomorpha; Acridoidea; Acrididae; Schistocerca.
 NC NBI_Taxid=7010;

RA [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=91092263; PubMed=1702381;
 RA Marshall J., Buckingham S.D., Shingai R., Lunt G.G., Goosey M.W.,
 RA Darlison M.G., Sattelle D.B., Barnard E.A.;
 RT "Sequence and functional expression of a single alpha subunit of an
 insect nicotinic acetylcholine receptor.";
 RL EMBO J. 9:4391-4398(1990).
 CC -1- FUNCTION: AFTER BINDING ACETYLCHOLINE, THE ACHR RESPONDS BY AN
 CC EXTENSIVE CHANGE IN CONFORMATION THAT AFFECTS ALL SUBUNITS AND
 CC LEADS TO OPENING OF AN ION-CONDUCTING CHANNEL ACROSS THE PLASMA
 CC MEMBRANE.
 CC -1- SUBCELLULAR LOCATION: Integral membrane protein.
 CC -1- SIMILARITY: BELONGS TO THE LIGAND-GATED IONIC CHANNELS FAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL: X55439; CAA39081.1; -.
 DR PIR: S12359; S12359.
 DR InterPro: IPR000188; GABAA_receptor.
 DR InterPro: IPR001175; Neur_channel.
 DR Pfam: PF02931; Neur_chan_LBD; 1.
 DR Pfam: PF02932; Neur_chan_memb; 1.
 DR PRINTS: PR00252; NRIONCHANNEL.
 DR PROSITE: PS00236; NEUROTR-ION_CHANNEL; 1.
 KW Receptor; Postsynaptic membrane; Ionic channel; Glycoprotein; Signal;
 KW Transmembrane; Multigene family.
 FT SIGNAL 1 23
 FT CHAIN 24 557
 FT DOMAIN 24 244
 FT TRANSMEM 245 266
 FT TRANSMEM 274 294
 FT TRANSMEM 308 329
 FT TRANSMEM 330 500
 FT TRANSMEM 501 523
 FT DISULFID 151 165
 FT DISULFID 224 225
 FT CARBOHYD 47 47
 FT CARBOHYD 235 235
 FT DOMAIN 382 400
 FT DOMAIN 406 422
 SQ SEQUENCE 557 AA; 63026 MW; 168389C887DFDF3E CRC64;

Query Match 17.6%; Score 407; DB 1; Length 557;
 Best Local Similarity 23.3%; Pred. No. 3.4e-25;
 Matches 128; Conservative 91; Mismatches 176; Indels 154; Gaps 19;

QY 17 PLLDMLMGLMGLTSGVATATADPKMANVKAIDRLHAGLFTNYSDDVQPGPTN 76
 DB 7 PMLL-----LLDLLLHNP-AAANPDAK-----RLYDGLLSNRYRLLRPSNNNTD- 51
 QY 77 VSELMVVYIIDDELNCK---LTTCHWLNRWDEERWQPSQYDNTQTLLKSSEVWTP 133
 DB 52 VLVIKGLRLSQLIDNLKDKDQITLTVNMLEHMDKFRMPDAEYGVTELYVPEHMLP 111
 QY 134 QITLFNDEGGIAME--TQVTLSDHGFRMPFPAVYTAACELNMLNMPHDKOSKIKGS 191
 DB 112 DIVLYNNADGEYVVTYTKAVLHHTGKVWTPPAIFKSSCEIDVRFPPDQCTEMKFGS 171
 QY 192 WGLKVVLPENGARGESLD-----HDD-----LVQSPMEIYDSRAHEVSQD 233
 DB 172 W-----TYGGDQDLKHNQKDDKKVKKGIDLRREYVEMDILIGVPAERREKY 221
 QY 234 Y-----YGYMEYTLTAQRSSMYTAVIYTPASCIIVILASAFWLPPHMGGEKIMINGL 287

```

DB 222 YPCSAEPDIPDFINTLRKTLFETVNLIVPCVGISYLVFLPAD-SGKIALCISI 280
OY 288 IIVIAAFMYFAQLPVLVSNNTLVVIFSTSLVSTLSTIVEVL---YLATGRKH-- 342
DB 261 LLSQTMFLLISLTPSTSLALPLGKYLFTWVLVGLSVTTIVLVNHYRKPSTHKMA 340
OY 343 -----RRLPEAL-----RKLLGH--LGTWLLSVSTTGESQAEK 376
DB 341 PWRKVFIRRLPKLLMRVPEQLDLASRLRHAHNSKLSAAAAAAYAAASSAASS 400
OY 377 TKEMDEHPPEADEQESSPLGINTHEVGC-----AKANQF----- 411
DB 401 PDSLHHHLHQ--HOHONHLQHLHQRPQGCNGLSHATNRFEGSGAGFGLPSVVGLDGS 458
OY 412 -----DMLLNTAVDRIFSVESLAF 432
DB 459 LSDVATRRKKYPEPELKAHNLVFIQNHMORODEPAEDQDMGVAMVLDRLFLMITIAS 518
OY 433 LI--LAIRC 439
DB 519 IYGTFAIIC 527

RESULT 13
ACHT_RAT STANDARD: PRT: 630 AA.
ID AC P09483: 035769:
AC 01-MAR-1989 (Rel. 10, Created)
BT 30-MAY-2000 (Rel. 39, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Neuronal acetylcholine receptor protein, alpha-4 chain precursor.
GN CHRM4 OR ACRA4.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Rattus.
OX NCBI_TaxId=10116:
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=SPRAGUE-DAWLEY; TISSUE=Hypothalamus, and Hippocampus;
RX MEDLINE=87159533; PubMed=3829125;
RA Goldman D.J., Denieris E.S., Luyten W., Kochhar A., Patrick J.,
RA Heinemann S.F.;
RT "Members of a nicotinic acetylcholine receptor gene family are
RT expressed in different regions of the mammalian central nervous
RT system."
RL Cell 46:965-973(1987).
RN [2]
RP REVISIONS.
RC STRAIN=SPRAGUE-DAWLEY; TISSUE=Brain;
RA Hatley M., Goldman D.J., Heinemann S.F.;
RL Submitted (DEC-1998) to the EMBL/GenBank/DBJ databases.
RN [3]
RP SEQUENCE FROM N.A.
RA Boulter J., Denieris E.S., Evans K., Heinemann S.F.;
RL Submitted (JUN-1997) to the EMBL/GenBank/DBJ databases.
RN [4]
RP PARTIAL SEQUENCE OF 31-47.
RX MEDLINE=87276531; PubMed=3609304;
RA Whiting P., Esch F., Shimazaki S., Lindstrom J.;
RT "Neuronal nicotinic acetylcholine receptor beta-subunit is coded for
RT by the cDNA clone alpha 4."
RL FEBS Lett. 219:459-463(1987).
CC -1- FUNCTION: AFTER BINDING ACETYLCHOLINE, THE ACHR RESPONDS BY AN
CC EXTENSIVE CHANGE IN CONFORMATION THAT AFFECTS ALL SUBUNITS AND
CC LEADS TO OPENING OF AN ION-CONDUCTING CHANNEL ACROSS THE PLASMA
CC MEMBRANE.
CC -1- SUBUNIT: NEURONAL ACHR SEEMS TO BE COMPOSED OF TWO DIFFERENT TYPES
CC OF SUBUNITS: ALPHA AND NON-ALPHA (BETA). ALPHA-4 SUBUNIT CAN BE
CC COMBINED TO BETA-2 OR BETA-4 TO GIVE RISE TO FUNCTIONAL RECEPTORS.
CC -1- SUBCELLULAR LOCATION: Integral membrane protein.
CC -1- ALTERNATIVE PRODUCTS: 2 ISOFORMS; ALPHA-4-1 (SHOWN HERE) AND
CC ALPHA-4-2; ARE PRODUCED BY ALTERNATIVE SPLICING.

```

```

CC -1- TISSUE SPECIFICITY: IN VARIOUS REGIONS OF THE CENTRAL NERVOUS
CC SYSTEM.
CC -1- SIMILARITY: BELONGS TO THE LIGAND-GATED IONIC CHANNELS FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see http://www.isb-sdb.ch/announce/
CC or send an email to license@isb-sdb.ch).
CC -----
CC EMBL: M15682; AAA41676.1; -.
CC EMBL: L31620; AAC97071.1; -.
CC EMBL: AF007212; AAB64439.1; -.
CC PIR: A26456; A26456.
CC InterPro: IPR000188; GABAA_receptor.
CC InterPro: IPR001175; Neur_chan.
CC Pfam: PF02931; Neur_chan_LBD; 1.
CC Pfam: PF02932; Neur_chan_memb; 1.
CC PRINTS: PR00252; NRIONCHANNEL.
CC PROSITE: PS00236; NEUROTR_ION_CHANNEL; 1.
CC PostSynaptic membrane; Ionic channel; Glycoprotein; Signal;
CC Transmembrane; Multigene family; Alternative splicing.
CC SIGNAL
CC CHAIN 1 31 630
CC NEURONAL ACETYLCHOLINE RECEPTOR PROTEIN,
CC ALPHA-4 CHAIN.
CC EXTRACELLULAR (POTENTIAL).
CC TRANSMEM 250 270
CC TRANSMEM 279 299
CC TRANSMEM 313 333
CC DOMAIN 334 604
CC TRANSMEM 605 625
CC DISULFD 163 177
CC DISULFD 227 228
CC CARBOHYD 59 59
CC CARBOHYD 109 109
CC CARBOHYD 176 176
CC VARSPIC 629 630
CC CONFLICT 1 4
CC CONFLICT 58 58
CC CONFLICT 196 196
CC CONFLICT 301 301
CC SEQUENCE 630 AA; 70192 MW; 1DFC9E1378E5417 CRC64;

Query Match 17.6%; Score 407; DB 1; Length 630;
Best Local Similarity 25.9%; Pred. No. 4e-25;
Matches 120; Conservative 84; Mismatches 176; Indels 84; Gaps 14;

OY 12 SGPGLPLLLQMLGMLGLTSPGATATADPKANVKAHLHAGLFTNYDSQVQYFO 71
DB 4 SGTGAPRPLLLPLLLLTGLTPASSHETRAHAEERLLKR---LFGYKMSRPV-- 57
OY 72 GRTNVSELEMYTY-----IDIDELNGKLTTHCHLNLRMDEDERVQWQSYDINTQT 124
DB 58 ---ANISDVVLVRFGLSTAQLLDVDEKQMTTNVWVKQEMDYKLRMDPGDEVNTSR 114
OY 125 LKSESEVFPQITLFFNGDEGGLAEQVTLSH--DGHFPMRPAYVATVACELNMLNMPHD 181
DB 115 ISELIMRDIYLVNNAADD-FAYVHLTAHLFYDGRQVMTPTAIYKSSCIDYVFFPPD 173
OY 182 KQSKLKIGSW--GLKVLPPENGARGESLDHDLVQSPWEIYDSRAHFVSODY--- 234
DB 174 QONCTMKFGSMTYDKAKIDL---VSMHSRYQDLDFWEGSEWIVDAVGYWTRREYCCA 229
OY 235 --YGMETVLTQNRSSMTAVIYTPASCIYVLLASAFPLPHMGKEKIMINGLLIYIA 292
DB 230 EIVPDITVAFTIRRLPLEFTTILIPCLLISCLYLVLFYVSEC-GEKVTLCISVLSLT 288
OY 293 AFLMYFAQLP-LVLSNTPLVVIFSTSLVSVSTIVEVLVYLATGRKH----- 342

```


GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OW protein - protein search, using sw model

Run on: July 18, 2002, 09:14:42 ; Search time 69.47 Seconds
(without alignments)
609.981 Million cell updates/sec

Title: US-09-732-680A-2

Sequence: 1 MTTTPKIKAPVSGPLPLLL.....RISFVSFLAFLILAIRCSV 441

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283138 segs, 96089334 residues

Total number of hits satisfying chosen parameters: 283138

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Database :
1: p1r1:*
2: p1r2:*
3: p1r3:*
4: p1r4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	444.5	19.2	528	1 ACCH2N	nicotinic acetylch
2	427	18.5	457	1 ACBOAI	nicotinic acetylch
3	424	18.3	511	2 A40110	nicotinic acetylch
4	418	18.1	627	2 JC4021	nicotinic acetylch
5	414.5	17.9	457	2 S13872	nicotinic acetylch
6	412.5	17.8	457	2 A24383	nicotinic acetylch
7	410	17.7	455	2 A55972	nicotinic acetylch
8	410	17.7	455	2 S11116	nicotinic acetylch
9	409.5	17.7	625	2 A26456	nicotinic acetylch
10	407.5	17.6	445	2 A19458	nicotinic acetylch
11	407.5	17.6	622	1 ACCH4N	nicotinic acetylch
12	407	17.6	557	2 S12359	nicotinic acetylch
13	406	17.6	494	2 T09289	nicotinic acetylch
14	405.5	17.5	457	1 ACHUAI	nicotinic acetylch
15	403.5	17.5	464	2 A33523	nicotinic acetylch
16	400.5	17.3	457	2 A28529	nicotinic acetylch
17	398.5	17.2	457	2 S08162	nicotinic acetylch
18	397	17.2	479	2 A55382	nicotinic acetylch
19	397	17.2	405	2 S07227	nicotinic acetylch
20	394.5	17.1	468	2 A38223	nicotinic acetylch
21	394.5	17.1	500	2 S12899	nicotinic acetylch
22	391.5	16.9	512	2 B37014	nicotinic acetylch
23	391	16.9	454	2 T24695	nicotinic acetylch
24	391	16.9	454	2 T24695	nicotinic acetylch
25	390.5	16.9	456	1 ACCH4N	nicotinic acetylch
26	390	16.9	493	2 S68587	nicotinic acetylch
27	389.5	16.9	461	2 T25671	nicotinic acetylch
28	389	16.8	461	2 T25671	hypothetical prote
29	389	16.8	523	2 T22910	hypothetical prote

30	384.5	16.6	501	2 A25338	nicotinic acetylch
31	383.5	16.6	499	2 A24572	nicotinic acetylch
32	378.5	16.4	461	1 ACRYA1	nicotinic acetylch
33	377.5	16.3	501	2 S04607	nicotinic acetylch
34	377	16.3	466	2 S16333	nicotinic acetylch
35	377	16.3	502	1 ACHUA7	nicotinic acetylch
36	377	16.3	502	2 A57175	nicotinic acetylch
37	376	16.3	462	2 S06893	nicotinic acetylch
38	374.5	16.2	542	2 T19862	hypothetical prote
39	372	16.1	501	2 S13873	nicotinic acetylch
40	372	16.1	502	2 G02259	alpha 7 neuronal n
41	371.5	16.1	493	1 ACRYB1	nicotinic acetylch
42	371	16.1	514	1 ACCHG1	nicotinic acetylch
43	369	16.0	423	2 T138056	nicotinic acetylch
44	369	16.0	502	2 T01378	nicotinic acetylch
45	368	15.9	503	2 A53956	nicotinic acetylch

ALIGNMENTS

RESULT 1

ACCH2N
nicotinic acetylcholine receptor alpha-2 chain precursor, neuronal - chicken
C:Species: Gallus gallus (chicken)
C:Date: 30-Sep-1991 #sequence_revision 30-Sep-1991 #text_change 21-Jul-2000
C:Accession: S00377
R:Net, P.: Oneyser, C.: Alliod, C.: Couturier, S.: Ballivet, M.
EMBO J. 7, 595-601, 1988
A:Title: Genes expressed in the brain define three distinct neuronal nicotinic acetylcholine receptor subunit families
A:Reference number: S00376; MUID:88283624
A:Accession: S00377
A:Molecule type: DNA
A:Residues: 1-528 <NER>
A:Cross-references: EMBL:X07339; NID:962792; PIDN:CAB59645.1; PID:96136914
A:Genetics:
A:Introns: 21/1; 73/3; 88/3; 125/2; 487/3
C:Superfamily: acetylcholine receptor
C:Keywords: glycoprotein; ion channel; neurotransmitter receptor; postsynaptic membrane
F:1-23/Domain: signal sequence #status predicted <SIG>
F:24-528/Product: nicotinic acetylcholine receptor alpha-2 chain #status predicted <M1>
F:241-264/Domain: transmembrane #status predicted <TM1>
F:272-290/Domain: transmembrane #status predicted <TM2>
F:306-327/Domain: transmembrane #status predicted <TM3>
F:502-520/Domain: transmembrane #status predicted <TM4>
F:54,104/Binding site: carbonylate (asn) (covalent) #status predicted
F:158-172,222-223/Disulfide bonds: #status predicted

Query Match 19.2% Score 444.5; DB 1; Length 528;

Best Local Similarity 27.7%; Pred. No. 1.7e-28; Matches 119; Conservative 82; Mismatches 14; Indels 81; Gaps 15;

QY	39	ATADPKANVAKLDRHAGLFTNYDSVQPVQGTPTNLSLMVVY-----IDDEL	91
DB	22	ATREQKQPHGFEDRLFKLFTGYNKRMSRP-----PNTSDVVIYKFGSLNQILDVDEK	76
QY	92	NKRLTHCWLNRMRDEEVWQPSQYDNTITTLASSEVWPTQITLFGNDEGLMAETOV	151
DB	77	NOMMTTNVWLKQWSDYKLRNMPDEDFDNTSIRVSEMIWIPDIYLYNNADGE-FAVTHM	135
QY	152	TLSSH---DGFRMPMPAVVAVYAVCELMNLPMDKSCXKLTGSW---GLKVVLPENGTAR	205
DB	136	TKAHLEFSNGKVKWVPAPATKSSCSIDVYFFPDQONCKMKFESWYDKAKIDL-----	188
QY	206	GSLDH---DDLVSPEWEIVDSRAHFVSQDY-----YGYMEYTLFAQRSSMYTAVI	255
DB	189	ENMEHVDLKDQWESGEAIIINAGRYNKKYDCDTEITPITFEYFARLPLFTIUL	247
QY	256	YRPASCTVTLASAFMLPPLPMGGEKIMGLIIVIAAFMLFAQLPLVLSNNTPLVYIF	315
DB	248	IPCLILISCTFVILVFLPSDC-GEKITLCISVLSLWFLTLITETIPSTSLVPIIGEX	306

```

QY 316 YSTSLYSLVSTVEVLVLYLNGKHKR-----RLPEALRKLHGLGTV-----LLL 363
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 307 LLEFTMIVFTLSITITVFLVNV-----HHRSPSTHMPHWSFFELGIPRMFLMKRPPLL 362
QY 364 SVSTSG-----ESQAEKTKEMDEHPHEADDESSPLGINTEV 403
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 363 PAECTGQYPPPGTRTSTRSCWLETVDVDDKWEKEEEEEEEEEEEKKAP-----SRV 417
QY 404 P--GAKANQ 410
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 418 PSGSGSQCTQ 426

RESULT 2
ACB0A1
nicotinic acetylcholine receptor alpha chain precursor - bovine
C:Species: Bos primigenius taurus (cattle)
C>Date: 18-Apr-1984 #sequence_revision 18-Apr-1984 #text_change 22-Jun-1999
C:Accession: A03169
R:Noda, M.; Furutani, Y.; Takahashi, H.; Toyosato, M.; Tanabe, T.; Shimizu, S.; Kikyoan
Nature 305, 818-823, 1983
A:Title: Cloning and sequence analysis of calf cDNA and human genomic DNA encoding alpha
A:Reference number: A03168; MUID:84039794
A:Accession: A03169
A:Molecule type: mRNA
A:Residues: 1-457 <NOD>
A:Cross-references: GB:X02509; NID:949; PIDN:CAA26345.1; PID:950
A:Note: four hydrophobic transmembrane segments are found in each of the four kinds of c
-318, and 429-447 in the alpha chain
C:Comment: The functional receptor molecule has two alpha chains and one each of the bet
C:Superfamily: acetylcholine receptor
C:Keywords: glycoprotein; ion channel; membrane protein; neurotransmitter receptor; post
F:1-20/Domain: signal sequence #status predicted <SIG>
F:21-457/Product: nicotinic acetylcholine receptor alpha chain #status predicted <MAT>
F:148-162/Disulfide bonds: #status predicted
F:161/Binding site: carbohydrate (Asn) (covalent) #status predicted

```

```

Query Match 18.38; Score 424; DB 2; Length 511;
Best Local Similarity 24.48; Pred. No. 8e-27;
Matches 126; Conservative 95; Mismatches 167; Indels 128; Gaps 18;

QY 20 LQMLMGLMGLTSVPATATADPRKANVKAIDRLHAGLFNTYSDVQVPFGQTPTVNSL 79
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 8 LQFTHLYLWCLLVP---AVLTQGGSHTHAEDRLFKHLFGYIRMARPV-----PMTSD 59
QY 80 EMVVTY-----IDDELNGKLTTHCWLNRWDEERWQPSOYDNIQTILKSSEVYT 132
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 60 VVIVREFGLSIQQLDVEDEKNQMTTNWKLKQEMNDYKLRMPAEFGVTSLRPSEMIWI 119
QY 133 PQLTFNGDESGMAEQVTLTSH---DGHFRMPRAYTATACELMNLNMPHDKQCKIKI 189
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 120 PDIVLYNNADGE-PAVTHMKAHLFETGYVHWVPALTKSCSIDYFFEPDQNCMKRF 178
QY 190 GSW---GLKVVLPENGARGESLDHDLVQSPEREIVDSRAHFVSODY-----YGYMEY 240
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 179 GSWYIDKAKILDEQ---MERTVDLKDYMSEGAIIINAGTYSKKYDDCAELIYRPVY 234
QY 241 TLTQRRSSMTAVIYTPASCIYVLAFAEPLPRHMGEEKIMINGLLIYIAFLMFAQ 300
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 235 YFVIRRLPLFETYLILIPCLLSCLFVLFVLPSEC-GEKTLTCLSVLSLTVFLITE 293
QY 301 LLPVLSNTPPVVVFYGSLSLYSVSTVEVLVLYLATGKHKR-----RLPEALRKLHG 355
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 294 IIPSTSLYIPLIGELFLTMIFVTLIVYFVLNV-----HHRSPSTHMPHWRVALLG 349
QY 356 HLGFWLLS-----VFSTGESQAETKEMDEHPHEADEQES 393
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 350 RVPRLMWNRPPLRPMELHGPSDLKSPSYHMLFTNMDAGEEETEE-----EEEDEN 403
QY 394 -----SPGLGNH-----TEVPAKANQF----- 411
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 404 ICWCAGLPDSSMGVLYGHGHLRAMPBEKTP--SQASEILSPQIKALEGVHYADRL 462
QY 412 -----DMALATAVDRISFVSFLAFLIA 436
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 463 RSEADSSVKEEMKVMVMDRI---FLWLFITVS 494

RESULT 4
JC4021
nicotinic acetylcholine receptor alpha-4 chain, neuronal - human
C:Species: Homo sapiens (man)
C>Date: 13-Jun-1995 #sequence_revision 14-Jul-1995 #text_change 20-Aug-1999
C:Accession: JC4021; S55471
R:Monteggia, L.M.; Gopalakrishnan, M.; Touma, E.; Idler, K.B.; Nash, N.; Arneric, S.P
Gene 155, 189-193, 1995
A:Title: Cloning and transient expression of genes encoding the human alpha 4 and bet
A:Reference number: JC4021; MUID:95237608

```


A:Accession: J04021
 A:Molecule type: mRNA
 A:Residues: 1-627 <MON>
 A:Cross-references: GB:L35901; NID:9755647; PIDN:AA64743.1; PID:9755648
 R:Manalali, A.; Remoundos, M.; Tzartos, S.
 submitted to the EMBL Data Library, May 1995
 A:Description: Molecular Cloning of human neuronal nicotinic acetylcholine receptor 4-1
 A:Reference number: S55471
 A:Accession: S55471
 A:Status: preliminary
 A:Molecule type: mRNA
 A:Residues: 26-627 <MAN>
 A:Cross-references: EMBL:X87629; NID:9854158; PIDN:CAA60959.1; PID:9854159
 C:Gene: GDB:CHRNA4
 A:Cross-references: GDB:128169; OMIM:118504
 A:Map position: 20q13.2-20q13.3
 C:Superfamily: acetylcholine receptor
 C:Keywords: ion channel; neurotransmitter receptor; postsynaptic membrane; transmembrane

Query Match 18.1%; Score 418; DB 2; Length 627;
 Best Local Similarity 29.8%; Pred. No. 3.3e-26;
 Matches 102; Conservative 67; Mismatches 139; Indels 34; Gaps 9;
 QY 11 VSGPGLPLDLMGLMGLTSPGATATADPKNANVKALDRLHAGLFTNYDSQVYF 70
 DB 3 LGPGAPRLPLDLLLGLTGLRASHVET-----RAHAERLLKLKLFSGNKKSRPV- 55
 QY 71 OGPTNVSLKVVY-----IDIDELNGKLTHCMLKRRDEERWQPSQYDNTIOI 123
 DB 56 ----ANISDVVLREGLSIAQLDVEDEKQMMTNNVWVQEWMDYKLRMDPADYEVNTSI 111
 QY 124 TLKSSVWTPQTLTFNGDGGIMAE--TQVTLSHDGHFRMPRAYTATCELMNLWMPD 181
 DB 112 RLPSELIMWPDVLVNNADGDAVTHLKAHLHGRQWTPRAYTKSSCIDYFFPPD 171
 QY 182 KOSCKLIKISW--GLKVLPRNGTAGESLDHDDLVSPEWEIVDSRAHFVSODY---- 234
 DB 172 QONCMKRESWYDRAKIDL-VNMHSR---VDQDPEWSEGEVIVDAVGTMTKRECCA 227
 QY 235 --YGMETYLTAORSSMTAVIYTPASCTIVLLASFWLPRHMGGEKIMINGLLIIVA 292
 DB 228 ELYPDITVAFVIRLRPLEFTINLLIPCLLISCTLVLFYLPSEC-GEKITLCSVLSLT 286
 QY 293 AFLMYFAQLPLVSNNTPLVVFYGSTLSLXSVSTIVEVLV 334
 DB 287 VFLILTEIPSTSLVPLIGELYLFTMTLFTVLSIVIVFL 328

RESULT 5
 S13872
 nicotinic acetylcholine receptor alpha chain precursor - rat
 C:Species: Rattus norvegicus (Norway rat)
 C:Date: 19-Mar-1997 #sequence_revision 19-Mar-1997 #text_change 20-Aug-1999
 C:Accession: S13872
 R:Witzemann, V.; Stein, E.; Barry, B.; Konno, T.; Koenen, M.; Kues, W.; Criado, M.; Hofme
 Eir, J. Biochem. 194, 437-448, 1990
 A:Title: Primary structure and functional expression of the alpha-, beta-, gamma-, delta
 A:Reference number: S13872; MUID:91099517
 A:Accession: S13872
 A:Status: preliminary
 A:Molecule type: mRNA
 A:Residues: 1-457 <MIT>
 A:Cross-references: EMBL:X74832; NID:9398831; PIDN:CAA52826.1; PID:9398832
 C:Superfamily: acetylcholine receptor

Query Match 17.9%; Score 414.5; DB 2; Length 457;
 Best Local Similarity 25.4%; Pred. No. 4.2e-26;
 Matches 118; Conservative 99; Mismatches 174; Indels 73; Gaps 19;
 QY 20 IQLMGLMLGLTSPGATATADPKNANVKALDRLHAGLFTNYDSQVYF---FOGTPTN 76

DB 1 MELTAVLLLLGLCSA-GTVLGSEHEH-----RLVAKLFKQYSSVRRVGDHRELVOYT 52
 QY 77 VSIEMVYITIDDELNGKLTHCHMLNRRDEERWQPSQYDNTIOITLKSEWVTPQT 136
 DB 53 VGIQL-IQILNDEVNQIATVNVRLKQWVDVNLKMPDQYGGVYKRIHPSEKIMRPDV 111
 QY 137 LFNQDGG--GLAEPOVTLSHDGHFRMPRAYTATCELMNLMPDKOSCKLKISW-- 192
 DB 112 LYNNADGFAIKFTKVLIDYGHITWTPPAIFKSCILITVHPPEBQNCMSKLGSTWY 171
 QY 193 --GLKVLDPENGARGESLDHDLVQSPWEIVDSRA--HFVSODYG-----YME--Y 240
 DB 172 DGSVAINPES-----DQDLSNFMESGEWVYKEARGKHNV---FYSCCPPTPIYDIT 223
 QY 241 TLTAQRSSMTAVIYTPASCTIVLLASFWLPRHMGGEKIMINGLLIIVAFLMYAQ 300
 DB 224 HEVMOQLPLFYFNVTIIPDLRFSLTSLVFYLPD--SGEKMILSISVLSLTFVFLVIVE 282
 QY 301 LPLVSNNTPLVVFYFSTSLVLSVSTIVEVLVYATGKHRR-----LPEAKRLHAG 355
 DB 283 LIPSTSSANPLIGKYLFTWVFVIASTITVIV---NTHRSBSTHIMPWVRKVFID 338
 QY 356 HGTWLLLSVFSTIGESQAER--TKEMDHFYEAEDEQESSPLG-----INHTVEPGA- 406
 DB 339 TPIPIMFSTMRKPSDKQEKRIETEDIDIS--DISGKRPMPGHSPLIKHPEVKAI 396
 QY 407 -----KANQ-----FDALIAVADRISEVFSFLAFI 434
 DB 397 EGVKVIATETMKSDQESNNASEMKYAVAMVDHLLGVFLVGLI 440

RESULT 6
 A24383
 nicotinic acetylcholine receptor alpha chain precursor - mouse
 C:Species: Mus musculus (house mouse)
 C:Date: 24-Jun-1987 #sequence_revision 24-Jun-1987 #text_change 15-Jun-1996
 C:Accession: A24383
 R:Isenberg, K.E.; Mudd, J.; Shah, V.; Merlie, J.P.
 Nucleic Acids Res. 14, 5111, 1986
 A:Reference number: A24383; MUID:86259081
 A:Accession: A24383
 A:Molecule type: mRNA
 A:Residues: 1-457 <ISE>
 C:Superfamily: acetylcholine receptor
 C:Keywords: glycoprotein; ion channel; neurotransmitter receptor; postsynaptic membra

Query Match 17.8%; Score 412.5; DB 2; Length 457;
 Best Local Similarity 25.8%; Pred. No. 6.1e-26;
 Matches 118; Conservative 96; Mismatches 177; Indels 73; Gaps 19;
 QY 26 MLMLGLTSPGATATADPKNANVKALDRLHAGLFTNYDSQVYF---FOGTPNVSLKVV 82
 DB 7 LLLGLCSA-GVLVGEHEH-----RLVAKLFEDYSSVVRVEDHREIVQVGLDL- 57
 QY 83 VYIIDDELNGKLTHCHMLNRRDEERWQPSQYDNTIOITLKSEWVTPQITFNGDE 142
 DB 58 IQLINDEVNQIATVNVRLKQWVDVNLKMPDQYGGVYKRIHPSEKIMRPDVVLYNNAD 117
 QY 143 G--GLMAETOVTLSHDGHFRMPRAYTATCELMNLMPDKOSCKLKISW--GLKV 196
 DB 118 GDPALVFKTKVLIDYGHITWTPPAIFKSCILITVHPPEBQNCMSKLGSTWY 177
 QY 197 VLPENGARGESLDHDLVQSPWEIVDSRA--HFVSODYG-----YME--YTLTAQR 246
 DB 178 INPES-----DQDLSNFMESGEWVYKEARGKHNV---FYSCCPPTPIYDITHFVQR 229
 QY 247 RSMSTAVIYTPASCTIVLLASFWLPRHMGGEKIMINGLLIIVAFLMYAQ 306
 DB 230 LPLVFTVNVITPCLLFSFTLSLTFYLPD--SGEKMILSISVLSLTFVFLVIVEIPST 288
 QY 307 NNTPLVVFYFSTSLVLSVSTIVEVLVYATGKHRR-----LPEALRKLLHGLHGLTWL 361

```

Db      289 SAVPLIGKMLFTVWVFVASTITIVYI---NTHNRSPSTHMEVWRKVFIDTIPNIM 344
      362 LLSVFTTGESQAEK---TKMEDHPYEAEDESSPLG-----INHEVGA----- 406
      345 FFSFMKRPSPNRKQKRIFTEDIDS--DISCKPGRPPGFSPLIKHEVSAIEGVYI 402
Qy      407 ----KANO-----FDMLATAVDRISFVSSFLAI 434
      403 AETMKSDQESNNAAEWKYVAMVDHLLGVFMLVCLL 440

```

RESULT 7

nicotinic acetylcholine receptor beta-3 chain precursor - chicken

C:Species: Gallus gallus (chicken)
C:Date: 03-Oct-1995 #sequence_revision 03-Oct-1995 #text_change 21-Jul-2000

C:Accession: A55972
R:Hernandez, M.C.; Erkman, L.; Matter-Sadzinska, L.; Roztocil, T.; Ballivet, M.; Matter, J. Biol. Chem. 270, 3224-3233, 1995

A:Title: Characterization of the nicotinic acetylcholine receptor beta3 gene. Its regula
A:Reference number: A55972; MUID:95155414

A:Accession: A55972
A:Status: preliminary

A:Molecule type: DNA
A:Residues: 1-455 <HER>

A:Cross-references: GB:X83739; NID:96165256; PIDN:CA59814.1; PID:96165257

C:Superfamily: acetylcholine receptor
C:Keywords: neurotransmitter receptor

Query Match 17.7%; Score 410; DB 2; Length 455;
Best Local Similarity 26.3%; Pred. No. 9.7e-26;
Matches 115; Conservative 85; Mismatches 167; Indels 70; Gaps 17;

```

Qy      52 DRLHAGFTNYSDVDPVFOGTPF-NVSLKMWVT-YIDIDELNCKLTTHCWLNRWDEE 109
      27 DALLRHLEFGYQKWKVRPVENSNDITKVLFGKISQLDVDEKQMLMTTNWLNKQEMMDHK 86
Qy      110 RVWQPSQDNTITQTLKSSEVWTPQITLF-NGD---EGGLMAETQVTLSHDGRFRMPRA 165
      87 LSNWPEEGGITAIRVSESLMPLDIYLFENADGRFEGSLM--TKAIVKYNQVQWMPRA 144
Qy      166 VYTAAYCELNLMPHDKOSCKLKIGSWG-----LKVVLPENGTARGESLDHDLVQSPW 220
      145 SYKSSCTMELTFPFDPQNCMKFGSWTYGDSWVDLLVD-----ENVDTKDFPDNGEW 198
Qy      221 EYDSRAHFVSQD---YGYMEYTLTAORSSMYTAVIYTPASCIYIALASAFMLPPHM 276
      199 EILNAKMKGNRKDGLSYFPVTVSYFLRRLPLFYTLFLIPLGLSFLVFLVFEYLPDSD- 257
Qy      277 GGEKIMINGLLIYIAFLMYFAQLPLVLSNMPLVVIFYSTSLYSVSTIVEVLVYL 336
      258 EGEKLSLSTVLSVLFVLVIEIIPSSSKVPLDIEYLFIMIFYTLSTIVVFIYV 317
Qy      337 ATGKHNR-----RPLEARKLHGLGTWLLSVFTTGESQAE-KT 377
      318 -----HHRSSATYHPMAVWRKLFQKLPRL--CMKGHVDRY-----SFSDEEKEETLKS 367
Qy      378 KEMDEHPYEAEDESSPLGINHTEVPG-----AKANOF-----DMALLATAVDRIS 424
      368 KLPKQKHQAKQDEKVIYAFLEKADSIYIRSHVKKDAFIROYVODMKFVAQVDRI- 426
Qy      425 FVSFSLAFLILAIRCSV 441
      427 ---FLMLFLVSVTGSV 440

```

RESULT 8

nicotinic acetylcholine receptor (nAChR) beta-3 chain precursor - chicken

C:Species: Gallus gallus (chicken)
C:Date: 07-May-1995 #sequence_revision 01-Sep-1995 #text_change 20-Aug-1999

C:Accession: S51116
R:Hernandez, M.C.; Erkman, L.; Matter-Sadzinska, L.; Roztocil, T.; Ballivet, M.; Matter, J. Submitted to the EMBL Data Library, January 1995
A:Description: Characterization of the nicotinic acetylcholine receptor beta-3 gene.
A:Reference number: S51116
A:Accession: S51116
A:Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-455 <HER>
A:Cross-references: EMBL:X83739; NID:9633182; PIDN:CA58707.1; PID:9633183
C:Superfamily: acetylcholine receptor
C:Keywords: neurotransmitter receptor

Query Match 17.7%; Score 410; DB 2; Length 455;
Best Local Similarity 26.3%; Pred. No. 9.7e-26;
Matches 115; Conservative 85; Mismatches 167; Indels 70; Gaps 17;

```

Qy      52 DRLHAGFTNYSDVDPVFOGTPF-NVSLKMWVT-YIDIDELNCKLTTHCWLNRWDEE 109
      27 DALLRHLEFGYQKWKVRPVENSNDITKVLFGKISQLDVDEKQMLMTTNWLNKQEMMDHK 86
Qy      110 RVWQPSQDNTITQTLKSSEVWTPQITLF-NGD---EGGLMAETQVTLSHDGRFRMPRA 165
      87 LSNWPEEGGITAIRVSESLMPLDIYLFENADGRFEGSLM--TKAIVKYNQVQWMPRA 144
Qy      166 VYTAAYCELNLMPHDKOSCKLKIGSWG-----LKVVLPENGTARGESLDHDLVQSPW 220
      145 SYKSSCTMELTFPFDPQNCMKFGSWTYGDSWVDLLVD-----ENVDTKDFPDNGEW 198
Qy      221 EYDSRAHFVSQD---YGYMEYTLTAORSSMYTAVIYTPASCIYIALASAFMLPPHM 276
      199 EILNAKMKGNRKDGLSYFPVTVSYFLRRLPLFYTLFLIPLGLSFLVFLVFEYLPDSD- 257
Qy      277 GGEKIMINGLLIYIAFLMYFAQLPLVLSNMPLVVIFYSTSLYSVSTIVEVLVYL 336
      258 EGEKLSLSTVLSVLFVLVIEIIPSSSKVPLDIEYLFIMIFYTLSTIVVFIYV 317
Qy      337 ATGKHNR-----RPLEARKLHGLGTWLLSVFTTGESQAE-KT 377
      318 -----HHRSSATYHPMAVWRKLFQKLPRL--CMKGHVDRY-----SFSDEEKEETLKS 367
Qy      378 KEMDEHPYEAEDESSPLGINHTEVPG-----AKANOF-----DMALLATAVDRIS 424
      368 KLPKQKHQAKQDEKVIYAFLEKADSIYIRSHVKKDAFIROYVODMKFVAQVDRI- 426
Qy      425 FVSFSLAFLILAIRCSV 441
      427 ---FLMLFLVSVTGSV 440

```

RESULT 9

nicotinic acetylcholine receptor alpha-4 chain, brain precursor - rat (fragment)

N:Alternate names: nicotinic acetylcholine receptor acetylcholine-binding chain; nAChR; N:Accession: A26456; MUID:87159533

C:Date: 31-Dec-1991 #sequence_revision 31-Dec-1991 #text_change 13-Nov-1998
R:Goldman, D.; Deneris, E.; Luyten, W.; Kocher, A.; Patrick, J.; Heinemann, S.

A:Title: Members of a nicotinic acetylcholine receptor gene family are expressed in d
A:Reference number: A26456; MUID:87159533

A:Accession: A26456
A:Molecule type: mRNA

A:Residues: 1-625 <GOLD>
A:Cross-references: EMBL:M15681; NID:9205637; PID:9205638

R:Whiting, P.; Esch, F.; Shimazaki, S.; Lindstrom, J.
FEBS Lett. 219, 459-463, 1987

A:Title: Neuronal nicotinic acetylcholine receptor beta-subunit is coded for by the c
A:Reference number: S02415; MUID:87276531

A:Accession: S02415
A:Molecule type: protein

A:Residues: 'X', 28-30, 'X', 32, 'X', 34-36, 'X', 38-40, 'X', 42-43 <WHIT>

C:Superfamily: acetylcholine receptor

C:Keywords: alternative splicing; glycoprotein; ion channel; neurotransmitter receptor;

F:1-26/Domain: signal sequence (fragment) #status predicted <SIG>

F:27-625/Product: nicotinic acetylcholine receptor alpha-4 chain #status experimental <M

Query Match 17.7%; Score 409.5; DB 2; Length 625;
Best Local Similarity 25.8%; Pred. No. 1.6e-25;
Matches 119; Conservative 84; Mismatches 176; Indels 83; Gaps 13;

QY 13 GCGCLPILLQMLGMLGLTSPGATATADPKNANKALDRHLACGLFTYDSDVQVFOG 72
DB 1 GCGAPPLPILLRLPGLLGLLPASSHETRAHAERLLKR---LFGSNKMSRPV--- 53
QY 73 TPTNLSLEWVYVY-----IDDELNGKLTTHCMNLNRPDEERVOXSDNITQITL 125
DB 54 --GNISDVIVLRPGLSIALIDVDEKNOMTTNVMVKQMDHDKKLMWDGDIENVTSTRI 111
QY 126 KSESWTPQITLFGNDEGGLMAETQVTLSH---DGHFRMPRAVYTAACELNMLNPHDK 182
DB 112 PSELIRPDIIVLYNNADG-FAVTHLTRKALFYDGRVQWTPATYKSSCSIDVTFEPFDQ 170
QY 183 QSCKLTIGSW---GLKVIPLPENGARGESLDHDDLVOSEFMEIVDSRAIFVSODY----- 234
DB 171 QNCTMKFGSWTDKAKIDL---VSIHSRVDDLDPEWSEGVNIVDAVGTNTRKCECAE 226
QY 235 -YGYMEYTLTAORSSMTAVIYTPASCIIVLLASFWLPRPHGGEKIMINGLIIVIAA 293
DB 227 TYPDIYATITRLPFTYINLIIFCLLSCLVLFVYSEC-GEKVLICISVLSLTV 285
QY 294 FLMTYQAQLLPVLSNNTPLVVFYSTSLYLSVSTIVEVLVLYLAKGNK----- 342
DB 286 PLLLTETIPSTSLVPLIGETLFTMIFVLTSLIVTVLVAN---NHSRPTHTPAMV 342
QY 343 -----RLPALKKLHNG-----HCTWLLSVFSTTGS 372
DB 343 RRVFIDIVRLFLMKRPVYKNCNRLISMKNAMAPFMEPEPGLSDICNOGLS 402
QY 373 QAETKEMDENHYEADDEQSSPLGINHTEVPGAKANOPDMA 414
DB 403 PAPTCNPDTAVETQPTCRSPPL-----EVPDLKTSEVKA 439

RESULT 10

I49458

acetylcholine receptor alpha-subunit - mouse (fragment)

C:Species: Mus musculus (house mouse)

C:Date: 02-Jul-1996 #sequence_revision 02-Jul-1996 #text_change 30-May-1997

C:Accession: I49458

R:Boulter, J.; Luyten, W.; Evans, K.; Mason, P.; Ballivet, M.; Goldman, D.; Stengelin, S

J. Neurosci. 5, 2545-2552, 1985

A:Title: Isolation of a clone coding for the alpha-subunit of a mouse acetylcholine rece

A:Reference number: I49458; MUID:85292055

A:Accession: I49458

A:Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: mRNA

A:Residues: 1-445 <RES>

A:Cross-references: GB:M17640; NID:9191601; PID:9191602

C:Superfamily: acetylcholine receptor

C:Keywords: neurotransmitter receptor

Query Match 17.6%; Score 407.5; DB 2; Length 445;
Best Local Similarity 26.0%; Pred. No. 1.5e-25;
Matches 112; Conservative 91; Mismatches 163; Indels 65; Gaps 17;

QY 53 RLHAGLFTYDSDVQV---FQGTPTNVSLEWVYIDDELNGKLTTHCMNLNRPDEE 109
DB 14 RLVAKLFEDYSSVVRPEVDHREIVQVTVGLQ-LQLINDEVQAVYTTNVRKQGVADN 72
QY 110 RWPQSQYDNIQITLKSESWTPQITLFGNDBG--GLMAEYQVTLSHGHRMPRAVY 167
DB 73 LKNMPPDYGVAKKIHIPSEKIMRPDVVLYNNADGFAIVKFTKVLVDYGHITWPPALF 132

QY 168 TAYCELNMLNMPHDKOSCKLTIGSW-----GLKVIPLPENGARGESLDHDDLVOSEFMEIV 223

DB 133 KSCETIYTHFPEDDNCNCKTCTWYDGSVAINEPES-----DQDLNFMESGEMVVK 187

QY 224 DSRN--HFVSODYG-----YME--YTLTAORSSMTAVIYTPASCIIVLLASFWLP 273

DB 188 EARGMKHNV---FYSCPTPTPYLDIYHFVMOQLPLFYIVNIIPLCLRSFLTSLVFYLP 244

QY 274 PHMGCKIMINGLIIVIAFLMTPAQLPVLNNTPLVVFYRSTSLYLSVSTIVEVLV 333

DB 245 TD-SGCKMPLTSLVLSLTVFLVFLVYELIPSTSAVPLGKXLMFTMVEVIAIITVIV 303

QY 334 LVLTATGKHRR-----LPALRKHLGHGTWLLSVFSTGSOAEK---TKRMDHPY 385

DB 304 I-----NTHHRSPTHTMPEVKKVFTDITPNIHFSTMKRPSDKOKERLFTEDIDIS-- 357

QY 386 EADDEQESSPLG-----INHTEVPGA-----KANQ-----FPMALLATAVDRI 423

DB 358 DISGRKPPPMGHPSPLIKHPKSAIEGVKYIAETKSDQESNNAEEMKYAVAMVDHI 417

QY 424 SFVPSFLALI 434

DB 418 LIGVFMVLCI 428

RESULT 11

ACCHAN

nicotinic acetylcholine receptor alpha-4 chain precursor, neuronal - chicken

C:Species: Gallus gallus (chicken)

C:Date: 30-Sep-1991 #sequence_revision 30-Sep-1991 #text_change 21-Jul-2000

C:Accession: S00379; A38756; A61186

R:Nef, P.; Oneyser, C.; Alliod, C.; Couturier, S.; Ballivet, M.

EMBO J. 7, 595-601, 1988

A:Title: Genes expressed in the brain define three distinct neuronal nicotinic acetyl

A:Reference number: S00376; MUID:88283624

A:Accession: S00379

A:Molecule type: mRNA

A:Residues: 1-622 <NEP>

A:Cross-references: GB:X07348; GB:Y00834; NID:962809; PIDN:CAA30285.1; PID:9871037

A:Accession: A38756

A:Molecule type: DNA

A:Residues: 21-622 <NE2>

A:Cross-references: EMBL:X07348

R:Whiting, P.J.; Schofer, R.; Conroy, W.G.; Gore, M.J.; Keyser, K.T.; Shimasaki, S.

Brain Res. Mol. Brain Res. 10, 61-70, 1991

A:Title: Expression of nicotinic acetylcholine receptor subtypes in brain and retina.

A:Reference number: A61186; MUID:91278688

A:Accession: A61186

A:Molecule type: protein

A:Residues: 24-29, 'XX', 32 <WHI>

C:Genetics:

A:introns: 21/1, 71/3; 86/3; 123/2; 581/3

C:Superfamily: acetylcholine receptor

C:Keywords: glycoprotein; ion channel; neurotransmitter receptor; postsynaptic membra

F:1-23/Domain: signal sequence #status predicted <SIG>

F:24-622/Product: nicotinic acetylcholine receptor alpha-4 chain #status experimental

F:239-262/Domain: transmembrane #status predicted <TM1>

F:270-288/Domain: transmembrane #status predicted <TM2>

F:304-325/Domain: transmembrane #status predicted <TM3>

F:596-614/Domain: transmembrane #status predicted <TM4>

F:52,102,388,472/Binding site: carbohydrate (asn) (covalent) #status predicted

F:156-170,220-221/Disulfide bonds: #status predicted

Query Match 17.6%; Score 407.5; DB 1; Length 622;
Best Local Similarity 29.4%; Pred. No. 2.4e-25;
Matches 96; Conservative 73; Mismatches 119; Indels 39; Gaps 10;

QY 50 ADRHLAGLFTYDSDVQVFGTPTNVSLEWVY-----IDDELNGKLTTHCMNLN 102
DB 31 ABERLLKLFSGYNKMSRPV-----ANISDVIVLRPGLSIALIDVDEKNOMTTNVMVK 85

OY 435 LAI 437
 Db 477 LCV 479

RESULT 14

ACHU1

nicotinic acetylcholine receptor alpha-1 chain precursor, muscle - human
 C:Species: Homo sapiens (man)

C:Date: 18-Apr-1984 #sequence_revision 18-Apr-1984 #text_change 22-Jun-1999

C:Accession: A03168; S00238; A27591

R:Noda, M.; Fututani, Y.; Takahashi, H.; Toyosato, M.; Tanabe, T.; Shimizu, S.; Kikuyotan
 Nature 305, 818-823, 1983

A:Title: Cloning and sequence analysis of calf cDNA and human genomic DNA encoding alpha
 A:Reference number: A03168; MUID:84039794

A:Accession: A03168

A:Molecule type: DNA

A:Residues: 1-457 <NOD>

A:Cross-references: GB:X02502; NID:g28391; PIDN:CAA26344.1; PID:g669153

R:Schoeffer, R.; Luther, M.; Lindstrom, J.
 FEBS Lett. 226, 235-240, 1988

A:Title: The human medulloblastoma cell line TE671 expresses a muscle-like acetylcholine

A:Reference number: S00238; MUID:88112190

A:Accession: S00238

A:Molecule type: mRNA

A:Residues: 1-457 <SCH>

A:Cross-references: EMBL:Y00762; NID:g28308; PIDN:CAA68731.1; PID:g28309

R:Hohlfeld, R.; Toyka, K.V.; Miner, L.L.; Walgrave, S.L.; Conti-Tronconi, B.M.
 J. Clin. Invest. 81, 657-660, 1988

A:Title: Amphipathic segment of the nicotinic receptor alpha subunit contains epitopes

A:Reference number: A27591; MUID:88139764

A:Contents: annotation

C:Genetics:

A:Gene: GDB:CHRNA1; CHRNA

A:Cross-references: GDB:120586; OMIM:100690

A:Map position: 2q24-q32

A:Intons: 15/1; 63/3; 78/3; 115/2; 180/3; 260/1; 334/3; 414/3

C:Complex: the functional receptor molecule is a heteropentamer with two alpha chains and

C:Superfamily: the acetylcholine receptor

C:Keywords: alternative splicing; glycoprotein; heteropentamer; ion channel; neurotransm

F:1-20/Domain: signal sequence #status predicted <SIG>

F:21-457/Product: nicotinic acetylcholine receptor alpha chain #status predicted <MAT>

F:231-257/Domain: transmembrane #status predicted <TM1>

F:263-281/Domain: transmembrane #status predicted <TM2>

F:297-318/Domain: transmembrane #status predicted <TM3>

F:429-447/Domain: transmembrane #status predicted <TM4>

F:148-162/Disulfide bonds: #status predicted

F:161/Binding site: carbohydrate (Asn) (covalent) #status predicted

Query Match 17.5%; Score 405.5; DB 1; Length 457;

Best Local Similarity 26.9%; Pred. No. 2.3e-25;

Matches 116; Conservative 84; Mismatches 165; Indels 67; Gaps 16;

OY 53 RLHAGLFTYDSDYDPV---FOGPTNVSLSEMYVTTIDIDELNGKLTTHCWLNRDERE 109

Db 26 RLVAKFKFYSSVVRVEHROVETVGLQ-IGLINDEVQVITTVNRKQKQVNDN 84

OY 110 RWQPSQYNTQITLTAKSEWTPQTITLNGDEG--GLMAETQVTLSDGHRMPPAVY 167

Db 85 LKWNPDYGVKVKIHIPSEKIMPRDLVLYNNADGFAIVKFTVLLQYTGHTWTPPAIF 144

OY 168 TAYCELMNLNMPHDKSCSLKIGSW---GLKVVLPENGTARGESLDHDDVOSPMEIY 223

Db 145 KSTCELTIVHFPPEQNCMKLGWTYDGSVAINPES---DQDLSNFMESGSGVWK 199

OY 224 DSRHAFVSODY-----GYMEYTLTAORSSMYTAVIYTPASCIYIALSAFMPPHM 276

Db 200 ESKGKHSVYISCCPTPLIDITVHVMQRLPYFTVNVIIIPCLLSFLGLVFLYPTD- 258

OY 277 GGEKIMINGLLIIVIAFLMFAQLLPVLSNNTPLVIVIFYSTLSLVSSTIVEVLVL 336

Db 259 SGEKMTLSISVLLSTLFTVLVIVELIPSTISAVPLIGKYMFLTMVIVIASIITVIV-- 316

OY 337 ATGKHRR-----LPEALRKLHGLTGLLSVFTGSGEQAETKREMDHPY-BEAD 390

Db 317 --NTHHRSPSTHVMVWKVF---IDTIPNMFSTMKRPSREK---ODKRTFEDIDI 368

OY 391 QESS-----PLG-----INTEVPGA-----KANQ-----FDMLLATAVDR 422

Db 369 SDISKRGPPGPPGPHSFLKHPVKSAIGIKYIAETMKSDDSNNAAEWKYVAAMVMDH 428

OY 423 ISVFSFLAFLI 434

Db 429 ILGVMVLCII 440

RESULT 15

A33523

nicotinic acetylcholine receptor beta-3 chain precursor - rat

C:Species: Rattus norvegicus (Norway rat)

C:Date: 29-Jan-1990 #sequence_revision 29-Jan-1990 #text_change 21-Jul-2000

C:Accession: A33523

R:Denneris, E.S.; Boulter, J.; Swanson, L.W.; Patrick, J.; Heinemann, S.

J. Biol. Chem. 264, 6268-6272, 1989

A:Title: Beta-3: a new member of nicotinic acetylcholine receptor gene family is expr

A:Reference number: A33523; MUID:89197926

A:Accession: A33523

A:Status: preliminary

A:Molecule type: mRNA

A:Residues: 1-464 <DEN>

A:Cross-references: GB:J04636; NID:g3298521; PIDN:AAC28887.1; PID:g205614

C:Superfamily: acetylcholine receptor

C:Keywords: neurotransmitter receptor; transmembrane protein

Query Match 17.5%; Score 403.5; DB 2; Length 464;

Best Local Similarity 24.9%; Pred. No. 3.4e-25;

Matches 121; Conservative 81; Mismatches 178; Indels 105; Gaps 17;

OY 18 LLLQMLMGMLNGLTGVSVPATATADPKNNVVALDRLHAGLTNTVSDYDPVFOGPTNV 77

Db 9 LVLSATLSGSWVTLVTRAGLSSVAENE-----DLRLHLLQGVOKKWRPVLN----- 55

OY 78 SLEMYVTV-----IDIDELNGKLTTHCWLNRDERVWQPSQYDNTQITLKSE 129

Db 56 SSDIIVGVGLKISQLVVDENKQMLTTVNWMLKQEWTDQKLRRNPREGYINSIKVPSDS 115

OY 130 VMTPTTLF-NGD---EGGLMAETQVTLSDGHRMPPAVIYATCELMLNMPHDKQSC 185

Db 116 LMLPDLVLEFNADGREGSLM--TKAIVKSSGTVSWTPASYSKSCMTMDVTFPPDRQNC 173

OY 186 KIKIGSW---GLKVVLPENGTARGESLDHDDVOSPMEIYDS-----RAHFVQDYX 235

Db 174 SKMFGSWYTDGTVNDL-----ILINENVDKRPFDNKGEMELLNKGMKGNRRGFTS--Y 226

OY 236 GMEYVTLTAORSSMYTAVIYTPASCIYIALSAFMPHMGGEKIMINGLLIIVIAFL 295

Db 227 PFTVTSFVLRRLPLFETFLIIPCLGSLFVLYVFLPBD-BEEKSLSTSVLSVLTVL 285

OY 296 MYFAQLLPVLSNNTPLVIVIFYSTLSLVSSTIVEVLVLATGKKRRLPEA--LRKLL 353

Db 286 LVTEELIPSSSKVYPLIGETLFLIMFVLSIVFVIVNHRSSSTYHPMAPVWRLLF 345

OY 354 HGLGTWLLSVSTGESAETKREMDHPYEADQESSPLGINHTFVPG----- 405

Db 346 LQRLPRMLCK-----DPMDRSPFDGESDPAVNG---RVSGKRKQTPAS 388

OY 406 -----AKANQ-----DMALLAVADRISFVSFLAFLIA 436

Db 389 DGERVLVAFLEKASESIRISRVKKEHFIISQVYQDMKRVAAQVLDRI-----FLMFLINS 444

OY 437 IRCV 441

Db 445 VLGS 449

Fri Jul 19 08:26:22 2002

us-09-732-680a-2.rpt

Page 8

Search completed: July 18, 2002, 10:19:32
Job time: 3890 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: July 18, 2002, 08:54:47 ; Search time 66.67 seconds
(without alignments)
161.567 Million cell updates/sec

Title: US-09-732-680A-2
Perfect score: 2311
Sequence: 1 MTTTPKIKAPVSGPLPLTL.....RISFVSFLAILAIRCSV 441

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 231628 seqs, 24425594 residues

Total number of hits satisfying chosen parameters: 231628

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :
1: Issued_Patents_AA.*
2: /cgn2_6/ptodata/2/1aa/5A_COMB.pep.*
3: /cgn2_6/ptodata/2/1aa/5B_COMB.pep.*
4: /cgn2_6/ptodata/2/1aa/6A_COMB.pep.*
5: /cgn2_6/ptodata/2/1aa/6B_COMB.pep.*
6: /cgn2_6/ptodata/2/1aa/Backfilest.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	423.5	18.3	629	US-08-464-258B-6	Sequence 6, Appli
2	423	18.3	511	US-08-464-258B-4	Sequence 4, Appli
3	422.5	18.3	629	US-08-278-635B-6	Sequence 6, Appli
4	422.5	18.3	629	US-08-471-961-6	Sequence 6, Appli
5	418.5	18.1	510	US-08-278-635B-4	Sequence 4, Appli
6	418.5	18.1	510	US-08-471-961-4	Sequence 4, Appli
7	416	18.0	529	US-08-496-855A-2	Sequence 2, Appli
8	414.5	17.9	457	US-08-278-635B-3	Sequence 3, Appli
9	414.5	17.9	457	US-08-471-961-3	Sequence 3, Appli
10	414	17.9	458	US-08-464-258B-3	Sequence 3, Appli
11	412.5	17.8	528	US-08-466-589-2	Sequence 2, Appli
12	412.5	17.8	528	US-08-700-636-2	Sequence 2, Appli
13	412.5	17.8	528	US-08-467-574-2	Sequence 2, Appli
14	412.5	17.8	528	US-09-217-345-2	Sequence 2, Appli
15	397	17.2	479	US-08-278-635B-2	Sequence 2, Appli
16	397	17.2	479	US-08-464-258B-2	Sequence 2, Appli
17	397	17.2	479	US-08-471-961-2	Sequence 2, Appli
18	384	16.6	627	US-08-466-589-6	Sequence 6, Appli
19	384	16.6	627	US-08-700-636-6	Sequence 6, Appli
20	384	16.6	627	US-08-467-574-6	Sequence 6, Appli
21	384	16.6	627	US-09-217-345-6	Sequence 6, Appli
22	381.5	16.5	449	5468481-1	Patent No. 5468481
23	381	16.5	502	US-08-278-635B-7	Sequence 7, Appli
24	381	16.5	502	US-08-464-258B-7	Sequence 7, Appli
25	381	16.5	502	US-08-471-961-7	Sequence 7, Appli
26	380.5	16.5	497	US-08-278-635B-5	Sequence 5, Appli
27	380.5	16.5	497	US-08-464-258B-5	Sequence 5, Appli

28	380.5	16.5	497	3	US-08-471-961-5	Sequence 5, Appli
29	373	16.1	502	2	US-08-466-589-8	Sequence 8, Appli
30	373	16.1	502	2	US-08-700-636-8	Sequence 8, Appli
31	373	16.1	502	3	US-08-467-574-8	Sequence 8, Appli
32	373	16.1	502	4	US-09-217-345-8	Sequence 8, Appli
33	371.5	16.1	449	4	US-08-462-351-2	Sequence 2, Appli
34	370.5	16.0	504	2	US-08-466-589-4	Sequence 4, Appli
35	370.5	16.0	504	2	US-08-700-636-4	Sequence 4, Appli
36	370.5	16.0	504	3	US-08-467-574-4	Sequence 4, Appli
37	370.5	16.0	504	4	US-09-217-345-4	Sequence 4, Appli
38	360.5	15.6	449	6	5194425-1	Patent No. 5194425
39	356	15.4	502	2	US-08-466-589-10	Sequence 10, Appli
40	356	15.4	502	2	US-08-700-636-10	Sequence 10, Appli
41	356	15.4	502	3	US-08-467-574-10	Sequence 10, Appli
42	356	15.4	502	4	US-09-217-345-10	Sequence 10, Appli
43	354.5	15.3	511	1	US-08-278-635B-8	Sequence 8, Appli
44	354.5	15.3	511	3	US-08-464-258B-8	Sequence 8, Appli
45	354.5	15.3	511	3	US-08-471-961-8	Sequence 8, Appli

ALIGNMENTS

```

RESULT 1
US-08-464-258B-6
; Sequence 6, Application US/08464258B
; Patent No. 6013766
;
GENERAL INFORMATION:
; APPLICANT: ELGOYHEN, ANA BELEN
; APPLICANT: JOHNSON, DAVID S.
; APPLICANT: BOULTER, JAMES R.
; APPLICANT: HEINEMANN, STEPHEN F.
; TITLE OF INVENTION: CLONING AND EXPRESSION OF A NOVEL
; NUMBER OF SEQUENCES: 8
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: GRAY CARY WARE & FREIDENRICH LLP
; STREET: 4365 EXECUTIVE DRIVE, SUITE 1600
; CITY: SAN DIEGO
; STATE: CALIFORNIA
; COUNTRY: USA
; ZIP: 92121
;
COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/464,258B
; FILING DATE: 06/05/95
; CLASSIFICATION: 530
; PRIORITY APPLICATION DATA:
; APPLICATION NUMBER: 08/278,635
; FILING DATE: 21-JUL-1994
; ATTORNEY/AGENT INFORMATION:
; NAME: REITER, STEPHEN E.
; REGISTRATION NUMBER: 31,192
; TELEPHONE: 619-677-1409
; TELEFAX: 619-677-1465
; INFORMATION FOR SEQ ID NO: 6:
; LENGTH: 629 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; US-08-464-258B-6

```

Query Match 18.3%; Score 423.5; DB 3; Length 629;
Best Local Similarity 26.1%; Pred. No. 5e-36;
Matches 121; Conservative 83; Mismatches 176; Indels 83; Gaps 13;

```

0Y      12  SGPGRPRLLOKMLGLMGLTSGVCAQTADPKANAKALDRLHAGFTNDSOVQFV 71
0Y      11  IIPGPIPLIOLIMLIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIII 71
Db      4  SGPGRPRLPLRLPLLLGLTGLLRASSIIEIRAAERLLKRL----LFSGKNKSRV-- 57
0Y      72  GTPPNVSLSENVYVY-----IDIDELNGKLTHHCWILMWREDERVWQPSQDNIQT 124
Db      58  ---GNISDVUPLVRFGLSLAQIOLIDVDEKNOAMTGVWVWKOEWNDYKLRMDPCDYENVT 114
0Y      125  LKSSVWVTPOTLTFNGDGGMACTQVYLSH---DGHFRMMRPVVYTAUCCINLMLNPHD 181
Db      115  IPSLIMRPDVLVNNMAGD--FATYHILTKAHLFYDGRVQWTPPAIVYSSSIDVTEPPFD 173
0Y      182  KQSCKLKIGSW---GLKVLVPENGTARGEISDHDLDVQSPEMEIVDSRAHFVSODY---- 234
Db      174  QONCKMKGSMVTYKAKIDL-----VSIHSRDQDLDFMSSGSMVYLDVAGVITNFKRYCCA 225
0Y      235  --YGYMEYTLTAQRSSMYTAIVITPASCIVIALSAFWLPRPHNGEKIMINGILLITVIA 292
Db      230  EIPYDTYAFIIRRLPLFYTTINLIIPCLLISCLTIVLFYVLPSEC--GKVYCCASSVLLST 288
0Y      293  AFIMYFQALLVLSNNPPLVYIFSTSLLYXSVSTIYEVLYLATYKHK----- 342
Db      289  VFLLITEIIPSTSLVPLJGEVLLFTVIEVTLISVITVEVLN---HHNSPRTHTPRAW 345
0Y      343  -----RRLPEALRKILHG-----HLGTWLLSVSTTGE 371
Db      346  VRVRFVLDVPLRLLEPMKRPSYVKDNCORRLIESMHKMANAPRPREPVEBPGILSDICNOGL 405
0Y      372  SOAEKTKEMDEHPYEADEQSSPLGINHTVEVPGAKANOEDMA 414
Db      406  SPAPFCNPTDTAVETOPTCRSPPL-----EVPDLKTSVEEKA 443

RESULT      2
US-08-464-258B-4
Sequence 4, Application US/08464258B
Patent No. 6013766
GENERAL INFORMATION:
APPLICANT: ELGOYHEN, ANA BELEN
APPLICANT: JOHNSON, DAVID S.
APPLICANT: BOULTER, JAMES R.
APPLICANT: HEINEMANN, STEPHEN F.
TITLE OF INVENTION: CLONING AND EXPRESSION OF A NOVEL
TITLE OF INVENTION: ACETYLCOLINE-GATED ION CHANNEL RECEPTOR
NUMBER OF SEQUENCES: 8
CORRESPONDENCE ADDRESS:
ADDRESSER: GRAY CARY WARE & FREIDENRICH LLP
STREET: 4365 EXECUTIVE DRIVE, SUITE 1600
CITY: SAN DIEGO
STATE: CALIFORNIA
COUNTRY: USA
ZIP: 92121
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/464,258B
FILING DATE: 06/05/95
CLASSIFICATION: 530
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/278,635
FILING DATE: 21-JUL-1994
ATTORNEY/AGENT INFORMATION:
NAME: REITER, STEPHEN E.
REGISTRATION NUMBER: 31,192
REFERENCE/DOCKET NUMBER: P41 9989
TELECOMMUNICATION INFORMATION:
TELEPHONE: 619-677-1403
TELEFAX: 619-677-1465
INFORMATION FOR SEQ ID NO: 4:

```

```

: SEQUENCE CHARACTERISTICS:
: LENGTH: 511 amino acids
: TYPE: amino acid
: STRANDEDNESS: single
: TOPOLOGY: linear
US-08-464-258B-4

Query Match      18.3%; Score 423; DB 3; Length 511;
Best Local Similarity 24.5%; Pred. No. 4e-36;
Matches 126; Conservative 94; Mismatches 167; Indels 128; Gaps 18;

QY 20 LQMLMGMLMGITSPVGATATADPKANAYKALDRLHAGLFTNYDSDVQPVFGPTPNVSL 79
   || : : : : || : : : : || : : : : || : : : : || : : : : ||
Db 8 LQFMTHLYLWCLLVP-----AVLQGSHTHADRLFKHLFGYNNMARPV-----PNTSD 59
   || : : : : || : : : : || : : : : || : : : : || : : : : ||

QY 80 EMVVIY-----IDIDELNGKLTHCHNLNKLWROBERWQPSOVDNITQILKSSEVWT 132
   || : : : : || : : : : || : : : : || : : : : || : : : : ||
Db 60 VVIATFEGUSIAQLDIVDKRNMQMTTNVWLKQEMWNYKRLKMDAEGWVYSLKVPSEMTVI 119
   || : : : : || : : : : || : : : : || : : : : || : : : : ||

QY 133 POLTFENDEGGIAMEQVTLSSH--DGHFRMPAPVATYACELNMLNMPHDKQSCKLKI 189
   || : : : : || : : : : || : : : : || : : : : || : : : : ||
Db 120 PDIVLYNNADDE-FAVTHMTKAHLFETGTGVHVVPATIKSSSIDIVTFEPFQDQCKMKF 178
   || : : : : || : : : : || : : : : || : : : : || : : : : ||

QY 190 GSW--GLKVVLPENGTRAGESLDHDDLQSGPEWELVDSRAHFVSODY-----YGYMEY 240
   || : : : : || : : : : || : : : : || : : : : || : : : : ||
Db 179 GSWIYDKARKIDLEQ---MERTVLDLKDWESGEEMAIINATGYNNSKKYDCCAEIYPDVAY 234
   || : : : : || : : : : || : : : : || : : : : || : : : : ||

QY 241 TLTQGRSSMTAVIYTPFASCIIVIALSAFMLPRHMGGEKIMINGLIIIVIAFLMYAQ 300
   || : : : : || : : : : || : : : : || : : : : || : : : : ||
Db 235 YFVIRRLPFTITNLIIPLCLLSCLTVLYFLIPSEC-GEKITLCISVLISLIVFLILLITE 293
   || : : : : || : : : : || : : : : || : : : : || : : : : ||

QY 301 LLPVLNNTPLVIFVYSTSLYLVSSTIVENLVLYLATGKNHR-----RLPEALRKLHG 355
   || : : : : || : : : : || : : : : || : : : : || : : : : ||
Db 294 IIPSTSLVPLIGEXYLLFTMIFVTLIVTVFLNV---HHRSPSTHMPMWVRVALLG 349
   || : : : : || : : : : || : : : : || : : : : || : : : : ||

QY 356 HLGTFLLLS-----VFSTGESQAKTKTEMDENHPRLEADEQES 393
   || : : : : || : : : : || : : : : || : : : : || : : : : ||
Db 350 RVPRLMNMNRPLPPNELHGSPLKLSPSYHMLETMMDAGERETE-----EEBEDEN 403
   || : : : : || : : : : || : : : : || : : : : || : : : : ||

QY 394 -----SPGLIHH-----TEVPKAKANOF----- 411
   || : : : : || : : : : || : : : : || : : : : || : : : : ||
Db 404 ICVCAGLEPDSMGVLYXGGLHRLAMEBETKRP-SQASILLSPDIOKALAEVHYIADRL 462
   || : : : : || : : : : || : : : : || : : : : || : : : : ||

QY 412 -----DWALLATAVDRISVSSSLAFLIL 435
   || : : : : || : : : : || : : : : || : : : : || : : : : ||
Db 463 RSEDADSSVKEDWKYVAMVDRI---FLMWELIIV 493
   || : : : : || : : : : || : : : : || : : : : || : : : : ||

RESULT 3
US-08-278-635B-6
: Sequence 6, Application US/08278635B
: Patent No. 5683912
: GENERAL INFORMATION:
: APPLICANT: ELGOYHEN, ANA BELEN
: APPLICANT: JOHNSON, DAVID S.
: APPLICANT: BOULDER, JAMES R.
: APPLICANT: HEINDMANN, STEPHEN F.
: TITLE OF INVENTION: CLONING AND EXPRESSION OF A NOVEL
: TITLE OF INVENTION: ACETYLCHOLINE-GATED ION CHANNEL RECEPTOR
: NUMBER OF SEQUENCES: 8
: CORRESPONDENCE ADDRESS:
: ADDRESSEE: GRAY CARY WARE & FREIDENRICH
: STREET: 4365 EXECUTIVE DRIVE, SUITE 1600
: CITY: SAN DIEGO
: STATE: CALIFORNIA
: COUNTRY: USA
: ZIP: 92121
: COMPUTER READABLE FORM:
: MEDIUM TYPE: Floppy disk
: COMPUTER: IBM PC compatible
: OPERATING SYSTEM: PC-DOS/MS-DOS
: SOFTWARE: PatentIn Release #1.0, Version #1.25

```



```

? APPLICANT: ELGOYHEN, ANA BELEN
? APPLICANT: JOHNSON, DAVID S.
? APPLICANT: BOULTER, JAMES R.
? APPLICANT: HEINEMANN, STEPHEN F.
? TITLE OF INVENTION: CLONING AND EXPRESSION OF A NOVEL
? TITLE OF INVENTION: ACETYLCHOLINE-GATED ION CHANNEL RECEPTOR
? NUMBER OF SEQUENCES: 8
? CORRESPONDENCE ADDRESS:
? ADDRESSEE: GRAY CARY WARE & FREIDENRICH
? STREET: 4365 EXECUTIVE DRIVE, SUITE 1600
? CITY: SAN DIEGO
? STATE: CALIFORNIA
? COUNTRY: USA
? ZIP: 92121
? COMPUTER READABLE FORM:
? MEDIUM TYPE: Floppy disk
? COMPUTER: IBM PC compatible
? OPERATING SYSTEM: PC-DOS/MS-DOS
? SOFTWARE: PatentIn Release #1.0, Version #1.25
? CURRENT APPLICATION DATA:
? APPLICATION NUMBER: US/08/278,635B
? FILING DATE: 21-JUL-1994
? CLASSIFICATION: 435
? ATTORNEY/AGENT INFORMATION:
? NAME: REITER, STEPHEN E.
? REGISTRATION NUMBER: 31,192
? REFERENCE/DOCKET NUMBER: P41 9771
? TELECOMMUNICATION INFORMATION:
? TELEPHONE: 619-677-1409
? TELEFAX: 619-677-1465
? INFORMATION FOR SEQ ID NO: 4:
? SEQUENCE CHARACTERISTICS:
? LENGTH: 510 amino acids
? TYPE: amino acid
? TOPOLOGY: linear
? MOLECULE TYPE: protein
? US-08-278-635B-4

```

```

Query Match 18.1%; Score 418.5; DB 1; Length 510;
Best Local Similarity 25.0%; Pred. No. 1.2e-35;
Matches 127; Conservative 88; Mismatches 178; Indels 115; Gaps 18;

```

```

QY 20 LQMLMGLMGLGTSVPGATATADPKNANVKALDRHLGGLFTNYDSVDPVFGQPTNVSL 79
DB 8 LQFWTHLYLWCLLVP---AVLTQGGSHTHAEDRLFKHLFGYNNRMAPV-----PMTSD 59
QY 80 EMVVTY-----IDIDELNGKLTTHCMNLNRDEERWVQPSQYDNTITQTLKSSEVWT 132
DB 60 VVIVRFGLSIAQLIDVDEKQNMNTNWMLKQEMNDYVNRMDPAEFGVTSLRVSEMIWI 119
QY 133 PQTLLFNGDEGGLMAETQVTLSH---DGHFRMPRAVYTYACELNMLNMPHDKOSCKLKI 189
DB 120 PDIYLVNNAAGE-FAVYHMTKAHLFFGTGVHWVPAIYKSSCSIDYFFPFDQNCMKKF 178
QY 190 GSW---GLKVVLPENGARGESLDHDLVOSPEWEIYDSRAHFVSQDY-----YGYMEY 240
DB 179 GSWTYDRAKIDLEO---MERTYDLKDYWESGEWAIINATGYNSKYDCCAEIYDPVY 234
QY 241 TLTRORSSMTAVIYTPASIVILASAPFLPRHMGSEKIMINGLLITVIAALMFAQ 300
DB 235 YFVIRRLPLFTYINLLIPCLISCLYLVAFYLPSEC-GEKITYLCISVLSLTVFLLLITE 293
QY 301 LLEPLNSNTPLVIVFYSTSLYLSTIVEYLVYLATGKHK-----RLPEALRKLHIG 355
DB 294 IIFSTSLVPIPLIGELYLEFTMIFVTLSTIYIVFLNV---HHRSPSHNMNMWVRVALLG 349
QY 356 HLGTVLL---LSVFTGESQAEKT-----KEMDEHPYEADQEE-----392
DB 350 RVPMLMNNRPLRPMELHSGPDLKLSPSYHLEFTNMDAGEEREETEEDENICVAGL 409
QY 393 -SSPLGINH-----TEVPGAKANQF-----411

```

```

DB 410 PDSSMGVLYCHGHLHRAPEETKP-SQASEILLSPOIKALEGVHYIADRLSEADS 468
QY 412 ----DMALLATAVDRISEVFSFLAILI 435
DB 469 SVKEDMKYVNAVVDRI-----FLMIFITV 492

```

```

RESULT 6
US-08-471-961-4
? Sequence 4, Application US/08471961
? Patent No. 6100046
? GENERAL INFORMATION:
? APPLICANT: ELGOYHEN, ANA BELEN
? APPLICANT: JOHNSON, DAVID S.
? APPLICANT: BOULTER, JAMES R.
? APPLICANT: HEINEMANN, STEPHEN F.
? TITLE OF INVENTION: CLONING AND EXPRESSION OF A NOVEL
? TITLE OF INVENTION: ACETYLCHOLINE-GATED ION CHANNEL RECEPTOR
? NUMBER OF SEQUENCES: 8
? CORRESPONDENCE ADDRESS:
? ADDRESSEE: GRAY CARY WARE & FREIDENRICH
? STREET: 4365 EXECUTIVE DRIVE, SUITE 1600
? CITY: SAN DIEGO
? STATE: CALIFORNIA
? COUNTRY: USA
? ZIP: 92121
? COMPUTER READABLE FORM:
? MEDIUM TYPE: Floppy disk
? COMPUTER: IBM PC compatible
? OPERATING SYSTEM: PC-DOS/MS-DOS
? SOFTWARE: PatentIn Release #1.0, Version #1.25
? CURRENT APPLICATION DATA:
? APPLICATION NUMBER: US/08/471,961
? FILING DATE:
? CLASSIFICATION:
? PRIOR APPLICATION DATA:
? APPLICATION NUMBER: US 08/278,635
? FILING DATE: 21-JUL-1994
? ATTORNEY/AGENT INFORMATION:
? NAME: REITER, STEPHEN E.
? REGISTRATION NUMBER: 31,192
? REFERENCE/DOCKET NUMBER: P41 9771
? TELECOMMUNICATION INFORMATION:
? TELEPHONE: 619-677-1409
? TELEFAX: 619-677-1465
? INFORMATION FOR SEQ ID NO: 4:
? SEQUENCE CHARACTERISTICS:
? LENGTH: 510 amino acids
? TYPE: amino acid
? TOPOLOGY: linear
? MOLECULE TYPE: protein
? US-08-471-961-4

```

```

Query Match 18.1%; Score 418.5; DB 3; Length 510;
Best Local Similarity 25.0%; Pred. No. 1.2e-35;
Matches 127; Conservative 88; Mismatches 178; Indels 115; Gaps 18;

```

```

QY 20 LQMLMGLMGLGTSVPGATATADPKNANVKALDRHLGGLFTNYDSVDPVFGQPTNVSL 79
DB 8 LQFWTHLYLWCLLVP---AVLTQGGSHTHAEDRLFKHLFGYNNRMAPV-----PMTSD 59
QY 80 EMVVTY-----IDIDELNGKLTTHCMNLNRDEERWVQPSQYDNTITQTLKSSEVWT 132
DB 60 VVIVRFGLSIAQLIDVDEKQNMNTNWMLKQEMNDYVNRMDPAEFGVTSLRVSEMIWI 119
QY 133 PQTLLFNGDEGGLMAETQVTLSH---DGHFRMPRAVYTYACELNMLNMPHDKOSCKLKI 189
DB 120 PDIYLVNNAAGE-FAVYHMTKAHLFFGTGVHWVPAIYKSSCSIDYFFPFDQNCMKKF 178
QY 190 GSW---GLKVVLPENGARGESLDHDLVOSPEWEIYDSRAHFVSQDY-----YGYMEY 240
DB 179 GSWTYDRAKIDLEO---MERTYDLKDYWESGEWAIINATGYNSKYDCCAEIYDPVY 234

```


Query Match 17.9%; Score 414.5; DB 1; Length 457;
Best Local Similarity 25.8%; Pred. No. 2,7e-35;
Matches 118; Conservative 97; Mismatches 170; Indels 73; Gaps 19;

QY 26 MLMLGLTSPGATATADPKNANVAKDLRLAGLTNDSDVOPV---FOGPTNVSLMV 82
DB 7 LLLGLSSA-GLVIGSHEH-----RLVAKLFEDYSSVVRPEHREIVQVTVGLQL- 57
QY 83 VTYIDIDELNGKLTTHCMLNLRMRDEERWVQPSQYDNIQTLLKSEVWTPQITLFGNDE 142
DB 58 IQLINDEVNQIYTTNRLKQWVDYLNKNNPDYGVKKIHIPSEKIMRPDVLVNNAD 117
QY 143 G--GLMAETOVTLSDHGFRMPRPVAYTAYCELMLNMPHDKOSCKLIKGSW---GLKV 196
DB 118 GDEFAIVFTVLLDYTGTHITWTPPAIKSYCEIIVTHFPPEDEONCSMKLTGWTYDGSVA 177
QY 197 VLPENGTARGESLDHDLVQSPMEIYDSRA--HFVSQDYG-----YME--YTLAQR 246
DB 178 INPES-----DQPLSNFMSEGEWVKEARGWKHW---FYSCPTTPYDITVHFVQR 229
QY 247 RSSWYTAIVTTPASCIYILALSAFWLPPHMGKEKIMINGLLIYIAFLMYFAQLLPVLS 306
DB 230 LPLFTYVNIIPCLFSEFLSVLYLPTD--SGEKMTLSISVLSLYFLVLYELLPTS 288
QY 307 NMPPLVIVYSTLSLYSVSTIVEVLYLATGKHRR-----LPEALRKLHGHGLTWL 361
DB 289 SAVPLIGKYMFTWVFVIAIIITVIVI---NTHHRSPTTHIMEVWRKVFIDITPINM 344
QY 362 LLSVFTSTGSOAEK---TEMDEHPYEADDEOSSPLG-----INTEVPGA----- 406
DB 345 FFSMKRPSRDKOKRFTEDIDIS--DISGKPGPPMGHSPILIKHPEVKSALIEGVKYI 402
QY 407 -----KANQ-----FDWALLATAVDRISEVFSFLAPLI 434
DB 403 AETMKSDQESNNAEMKYYAMWMDHLLGVFMVLCI 440

RESULT 9
US-08-471-961-3
; Sequence 3, Application US/08471961
; Patent No. 6100046
; GENERAL INFORMATION:
; APPLICANT: ELGOYHEN, ANA BELEN
; APPLICANT: JOHNSON, DAVID S.
; APPLICANT: BOULTER, JAMES R.
; APPLICANT: HEINEMANN, STEPHEN F.
; TITLE OF INVENTION: CLONING AND EXPRESSION OF A NOVEL
; TITLE OF INVENTION: ACETYLCHOLINE-GATED ION CHANNEL RECEPTOR
; NUMBER OF SEQUENCES: 8
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: GRAY CARY WARE & FREIDENRICH
; STREET: 4365 EXECUTIVE DRIVE, SUITE 1600
; CITY: SAN DIEGO
; STATE: CALIFORNIA
; COUNTRY: USA
; ZIP: 92121
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: IBM PC compatible
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/471,961
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/278,635
; FILING DATE: 21-JUL-1994
; ATTORNEY/AGENT INFORMATION:
; NAME: REITER, STEPHEN E.
; REGISTRATION NUMBER: 31,192
; REFERENCE/DOCKET NUMBER: P41 9771

TELECOMMUNICATION INFORMATION:
; TELEPHONE: 619-677-1409
; TELEFAX: 619-677-1465
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 457 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-471-961-3

Query Match 17.9%; Score 414.5; DB 3; Length 457;
Best Local Similarity 25.8%; Pred. No. 2,7e-35;
Matches 118; Conservative 97; Mismatches 170; Indels 73; Gaps 19;

QY 26 MLMLGLTSPGATATADPKNANVAKDLRLAGLTNDSDVOPV---FOGPTNVSLMV 82
DB 7 LLLGLSSA-GLVIGSHEH-----RLVAKLFEDYSSVVRPEHREIVQVTVGLQL- 57
QY 83 VTYIDIDELNGKLTTHCMLNLRMRDEERWVQPSQYDNIQTLLKSEVWTPQITLFGNDE 142
DB 58 IQLINDEVNQIYTTNRLKQWVDYLNKNNPDYGVKKIHIPSEKIMRPDVLVNNAD 117
QY 143 G--GLMAETOVTLSDHGFRMPRPVAYTAYCELMLNMPHDKOSCKLIKGSW---GLKV 196
DB 118 GDEFAIVFTVLLDYTGTHITWTPPAIKSYCEIIVTHFPPEDEONCSMKLTGWTYDGSVA 177
QY 197 VLPENGTARGESLDHDLVQSPMEIYDSRA--HFVSQDYG-----YME--YTLAQR 246
DB 178 INPES-----DQPLSNFMSEGEWVKEARGWKHW---FYSCPTTPYDITVHFVQR 229
QY 247 RSSWYTAIVTTPASCIYILALSAFWLPPHMGKEKIMINGLLIYIAFLMYFAQLLPVLS 306
DB 230 LPLFTYVNIIPCLFSEFLSVLYLPTD--SGEKMTLSISVLSLYFLVLYELLPTS 288
QY 307 NMPPLVIVYSTLSLYSVSTIVEVLYLATGKHRR-----LPEALRKLHGHGLTWL 361
DB 289 SAVPLIGKYMFTWVFVIAIIITVIVI---NTHHRSPTTHIMEVWRKVFIDITPINM 344
QY 362 LLSVFTSTGSOAEK---TEMDEHPYEADDEOSSPLG-----INTEVPGA----- 406
DB 345 FFSMKRPSRDKOKRFTEDIDIS--DISGKPGPPMGHSPILIKHPEVKSALIEGVKYI 402
QY 407 -----KANQ-----FDWALLATAVDRISEVFSFLAPLI 434
DB 403 AETMKSDQESNNAEMKYYAMWMDHLLGVFMVLCI 440

RESULT 10
US-08-464-258B-3
; Sequence 3, Application US/08464258B
; Patent No. 6013766
; GENERAL INFORMATION:
; APPLICANT: ELGOYHEN, ANA BELEN
; APPLICANT: JOHNSON, DAVID S.
; APPLICANT: BOULTER, JAMES R.
; APPLICANT: HEINEMANN, STEPHEN F.
; TITLE OF INVENTION: CLONING AND EXPRESSION OF A NOVEL
; TITLE OF INVENTION: ACETYLCHOLINE-GATED ION CHANNEL RECEPTOR
; NUMBER OF SEQUENCES: 8
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: GRAY CARY WARE & FREIDENRICH LLP
; STREET: 4365 EXECUTIVE DRIVE, SUITE 1600
; CITY: SAN DIEGO
; STATE: CALIFORNIA
; COUNTRY: USA
; ZIP: 92121
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: IBM PC compatible
; SOFTWARE: PatentIn Release #1.0, Version #1.25

CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/08/464,258B
 FILING DATE: 06/05/95
 CLASSIFICATION: 530
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: 08/278,635
 FILING DATE: 21-JUL-1994
 ATTORNEY/AGENT INFORMATION:
 NAME: REITER, STEPHEN E.
 REGISTRATION NUMBER: 31,192
 REFERENCE/DOCKET NUMBER: P41 9989
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: 619-677-1409
 TELEFAX: 619-677-1465
 INFORMATION FOR SEQ ID NO: 3:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 458 amino acids
 TYPE: amino acid
 STRANDEDNESS: single
 TOPOLOGY: linear
 US-08-464-258B-3

Query Match 17.9%; Score 414; DB 3; Length 458;
 Best Local Similarity 25.7%; Pred. No. 3e-35;
 Matches 118; Conservative 97; Mismatches 170; Indels 74; Gaps 19;

QY 26 MLMLGTVPGATATADKNNANVADLRHAGLFNTYDSQVY---FQGTPTNVSLEMY 82
 DB 7 LLLGLISSA-GVLYGSEHET-----RLVAKLFEDYSSVVRDHRRIYQVTVGLD- 57
 QY 83 VTYIDIDELNGKLTTHCMNLKRMDEERYWQSOYDNITQITLKSEVWPQITLFNGDE 142
 DB 58 IOLINDEVNOQVITNVRNLKQOWVYNLKNRPDYGCVKIKIIPSEKIKRPVVLXNNAD 117
 QY 143 G--GMAETQVTLSDHGRFMRMPRAVYTAUCLENNLMPHDKQSKLKGSW---GLKV 196
 DB 118 GQFAIVKFTKVLDTGHTWTPRAIFKSYCEIIVTHPEFDQNSMKLGTVTQGSVVA 177
 QY 197 VLPENGTAGESLDHDLVQSPREMEIVDSRA--HFVSQDYG-----YME--YTLTAOR 246
 DB 178 INPE-----DQPDLSNMESEGENYIKERKGMHW--FYSCPTPTVLDITVHFVMQR 229
 QY 247 RSSMTAVIYVPPASCIIVIALSAFWLPRMGEKIMINGLLIIVIAFLMYRAQLLPVLS 306
 DB 230 LPLRYIVAVIIPCLFSEFLTSVFTLPD-SEKMTLISVLSTVFLVLVELIPSTIS 288
 QY 307 NNPPLVIFYSTSLYLSVSTIVEVLVLYLATGKHRR---LRPALRKLHGLGTLW 361
 DB 289 SAVPLIGKMYLFTVWFVLAIIIVYI---NTHHRSPSTHIMPEWAKVFDITIPNIM 344
 QY 362 LLSVSTGESQAEK---TKMDENHYEADQESSPLG-----INTEVGA----- 406
 DB 345 FFSYTKRRSRDQEKRIFFEDIDIS--DISGKRPGRPPMGFHSPLIKHPEVSALEGVY 402
 QY 407 -----KANO-----EDMALLATAVDRISEFSLAFIL 434
 DB 403 IAEYKMSDQESNNAABEKYVAMVMDHILGVFMVLCIL 441

RESULT 11
 US-08-466-589-2
 ; Sequence 2, Application US/08466589
 ; Patent No. 5837489
 ; GENERAL INFORMATION:
 ; APPLICANT: Elliot, Kathryn J.
 ; APPLICANT: Ellis, Steven B.
 ; APPLICANT: Harpold, Michael M.
 ; TITLE OF INVENTION: HUMAN NEURONAL NICOTINIC ACETYLCHOLINE
 ; NUMBER OF SEQUENCES: 12
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: Brown, Martin, Haller & McClaim

STREET: 1660 Union Street
 CITY: San Diego
 STATE: CA
 COUNTRY: USA
 ZIP: 92101-2926
 COMPUTER READABLE FORM:
 MEDIUM TYPE: Diskette
 COMPUTER: IBM Compatible
 OPERATING SYSTEM: DOS
 SOFTWARE: FASTSEQ Version 1.5
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/08/466,589
 FILING DATE: June 5, 1995
 CLASSIFICATION: 536
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: US 08/028,031
 FILING DATE: March 8, 1993
 ATTORNEY/AGENT INFORMATION:
 NAME: Seidman, Stephanie L.
 REGISTRATION NUMBER: 33,779
 REFERENCE/DOCKET NUMBER: 6362-9950
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: 619-238-0999
 TELEFAX: 619-238-0062
 TELEX:
 INFORMATION FOR SEQ ID NO: 2:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 528 amino acids
 TYPE: amino acid
 STRANDEDNESS: single
 TOPOLOGY: unknown
 MOLECULE TYPE: protein
 US-08-466-589-2

Query Match 17.8%; Score 412.5; DB 2; Length 528;
 Best Local Similarity 27.4%; Pred. No. 5.5e-35;
 Matches 108; Conservative 76; Mismatches 153; Indels 57; Gaps 12;

QY 13 GGPLLLLOML-MGLMLGTVPGATATADPKNA-----NKKALDR 53
 DB 2 GSCCVFLSTFTKLSLMLLITPAGGEAKRPPRAPGDLSPSPALAPQGSHTETEDR 61
 QY 54 LHAGLFTNVDSQVYFQGTPTNVSLEMYVY-----IDIDELNGKLTTHCMNLKRR 106
 DB 62 LFKHLFRGYNRRARVY-----PNTSDVYIVRFGLISIAQILDVDERNOMATTNWLKQENS 116
 QY 107 DEERYWQSOYDNITQITLKSEVWPQITLFNGDEGGLMAETQVTLSH---DGHFRMP 163
 DB 117 DYKLRMPADPFENITISLRVPSMIMIPDIVLYN--KXGEFAVTHMTKHLFSTGVHMYR 174
 QY 164 PAVYTAUCLENNLMPHDKQSKLKGSW---GLKVLPENGTAGESLDHDLVQSPBW 220
 DB 175 PAIYKSSCSIDVTPEFPDQNCMKFSGSWTYDKAKIDLEQ---MEQFVLDKYWESGHW 230
 QY 221 EIVDSRAHFVSQDY-----YGYMEYTLTAORSSMYAVIYTPASCIIVIALSAFWLBP 274
 DB 231 AIVNATGTYNSKKYDCCAEIIPDVYAVYIRKPLFTIINLLIIPCLISCLVLYVFLDS 290
 QY 275 HMGGEKIMINGLLIIVIAFLMYFAQLPVLNSNTPLVVIFYSTSLYLSVSTIVEVLV 334
 DB 291 DC-GERKITLCISVLSTVFEFLITIEIIPSTSLVYIPLIGEYLLFMIFVTLISVITFVL 349
 QY 335 YLATGKHK---RLPALRKLHGLGTLWLLS 364
 DB 350 NV---DHRSPSTHTMPHWKVALGCVPRMLIMN 380

RESULT 12
 US-08-700-636-2
 ; Sequence 2, Application US/08700636
 ; Patent No. 5910582
 ; GENERAL INFORMATION:

```

? APPLICANT: Elliott, Kathryn J.
? APPLICANT: Ellis, Steven B.
? APPLICANT: Harpold, Michael M.
? TITLE OF INVENTION: HUMAN NEURONAL NICOTINIC ACETYLCHOLINE
? TITLE OF INVENTION: RECEPTOR COMPOSITIONS AND METHODS EMPLOYING SAME
? NUMBER OF SEQUENCES: 12
? CORRESPONDENCE ADDRESS:
? ADDRESSEE: Pretty, Schroeder, Brueggemann & Clark
? STREET: 444 South Flower Street, Suite 2000
? CITY: Los Angeles
? STATE: CA
? COUNTRY: USA
? ZIP: 90071
? COMPUTER READABLE FORM:
? MEDIUM TYPE: Floppy disk
? COMPUTER: IBM PC compatible
? OPERATING SYSTEM: PC-DOS/MS-DOS
? SOFTWARE: Patentin Release #1.0, Version #1.25
? CURRENT APPLICATION DATA:
? APPLICATION NUMBER: US/08/700,636
? FILING DATE: 16-JUL-1996
? CLASSIFICATION: 536
? PRIOR APPLICATION DATA:
? APPLICATION NUMBER: US 08/028,031
? FILING DATE: 08-MAR-1993
? ATTORNEY/AGENT INFORMATION:
? NAME: Reiter, Stephen E.
? REGISTRATION NUMBER: 31,192
? REFERENCE/DOCKET NUMBER: P41 9368
? TELECOMMUNICATION INFORMATION:
? TELEPHONE: 619-546-4737
? TELEFAX: 619-546-9392
? INFORMATION FOR SEQ ID NO: 2:
? SEQUENCE CHARACTERISTICS:
? LENGTH: 528 amino acids
? TYPE: amino acid
? STRANDEDNESS: single
? TOPOLOGY: unknown
? MOLECULE TYPE: protein
? US-08-700-636-2

```

```

Query Match      17.8%; Score 412.5; DB 2; Length 528;
Best Local Similarity 27.4%; Pred. No. 5,5e-35;
Matches 108; Conservative 76; Mismatches 153; Indels 57; Gaps 12;

? 13 GPGPLPILQL-MGMILMGLTSVGATATADPKNA-----NVRALDR 53
? 2 GPCSPVFLSTKLSLWMLLTPAGGEAKRPPRAPGDDPLSSPSPTALPOGGSHTETEDR 61
? 54 LHAGLFTNYSDDVQPVQGTPTNVSLSEMYTY-----IDDELNGKLTTHGWLNRMR 106
? 62 LFNHLFRGYRMARPV-----PNTSDVIVIRFGISIAQLIDVDEKKNOMMTNWLKQEMS 116
? 107 DEERVWQPSQYDNITQITLTKSSSEVWTPQITLFGNDEGLMAETQVLSH---DGHFRMP 163
? 117 DYLRNMPADFGNITSLRVSEMIWIPDIYLN--KXGEFAVYHMTKAHLFSTGYVHWVP 174
? 164 PAYTAYACELNMLNMPHDKOSCKLIGSW---GLKVVLPENGTARGESLDHDDLVOGSPW 220
? 175 PAIYKSSCIDVYFFPPDOONCKMKFGSWTYDKAKIDLEO---MEQTVLDKDYWESGEW 230
? 221 EIVDSRAHFVSODY-----YGYMEYTLTAQRSSMYTAVIYTPASCIVILASAFWLP 274
? 231 AIYNATGYNSKKYDCCAEIYPDVYAFVIRLPLEFTINLILPCLLISCLTVLVEFLPS 290
? 275 HMGCKMIMGLLIYIAAFLMFAQLPLVLSNTPLVVIFYSTSLIYSSTIVEYVL 334
? 291 DC-GEKTLICISVLSLTVFLLEITELIPSTSVIPIGELYLLFTMTFVLSIVYEV 349
? 335 YLATGKHK---RLPLEALRKLHGHGTLWLLS 364
? 350 NV---DHRSPSTHTMPHWVRGALLGCVPRWLLMN 380

```

```

RESULT 13
US-08-467-574-2
? Sequence 2, Application US/08467574
? Patent No. 6022704
? GENERAL INFORMATION:
? APPLICANT: Elliott, Kathryn J.
? APPLICANT: Ellis, Steven B.
? APPLICANT: Harpold, Michael M.
? TITLE OF INVENTION: HUMAN NEURONAL NICOTINIC ACETYLCHOLINE
? TITLE OF INVENTION: RECEPTOR COMPOSITIONS AND METHODS EMPLOYING SAME
? NUMBER OF SEQUENCES: 12
? CORRESPONDENCE ADDRESS:
? ADDRESSEE: Brown, Martin, Haller & McClaim
? STREET: 1660 Union Street
? CITY: San Diego
? STATE: CA
? COUNTRY: USA
? ZIP: 92101-2926
? COMPUTER READABLE FORM:
? MEDIUM TYPE: Diskette
? OPERATING SYSTEM: DOS
? SOFTWARE: FastSeq Version 1.5
? CURRENT APPLICATION DATA:
? APPLICATION NUMBER: US/08/467,574
? FILING DATE: June 5, 1995
? CLASSIFICATION: 536
? PRIOR APPLICATION DATA:
? APPLICATION NUMBER: US 08/028,031
? FILING DATE: March 8, 1993
? ATTORNEY/AGENT INFORMATION:
? NAME: Seidman, Stephanie L.
? REGISTRATION NUMBER: 33,779
? REFERENCE/DOCKET NUMBER: 6362-9949
? TELECOMMUNICATION INFORMATION:
? TELEPHONE: 619-238-0999
? TELEFAX: 619-238-0062
? TELEX:
? INFORMATION FOR SEQ ID NO: 2:
? SEQUENCE CHARACTERISTICS:
? LENGTH: 528 amino acids
? TYPE: amino acid
? STRANDEDNESS: single
? TOPOLOGY: unknown
? MOLECULE TYPE: protein
? US-08-467-574-2

```

```

Query Match      17.8%; Score 412.5; DB 3; Length 528;
Best Local Similarity 27.4%; Pred. No. 5,5e-35;
Matches 108; Conservative 76; Mismatches 153; Indels 57; Gaps 12;

? 13 GPGPLPILQL-MGMILMGLTSVGATATADPKNA-----NVRALDR 53
? 2 GPCSPVFLSTKLSLWMLLTPAGGEAKRPPRAPGDDPLSSPSPTALPOGGSHTETEDR 61
? 54 LHAGLFTNYSDDVQPVQGTPTNVSLSEMYTY-----IDDELNGKLTTHGWLNRMR 106
? 62 LFNHLFRGYRMARPV-----PNTSDVIVIRFGISIAQLIDVDEKKNOMMTNWLKQEMS 116
? 107 DEERVWQPSQYDNITQITLTKSSSEVWTPQITLFGNDEGLMAETQVLSH---DGHFRMP 163
? 117 DYLRNMPADFGNITSLRVSEMIWIPDIYLN--KXGEFAVYHMTKAHLFSTGYVHWVP 174
? 164 PAYTAYACELNMLNMPHDKOSCKLIGSW---GLKVVLPENGTARGESLDHDDLVOGSPW 220
? 175 PAIYKSSCIDVYFFPPDOONCKMKFGSWTYDKAKIDLEO---MEQTVLDKDYWESGEW 230
? 221 EIVDSRAHFVSODY-----YGYMEYTLTAQRSSMYTAVIYTPASCIVILASAFWLP 274
? 231 AIYNATGYNSKKYDCCAEIYPDVYAFVIRLPLEFTINLILPCLLISCLTVLVEFLPS 290

```

```

0Y      54 LHAGLEFNYSDDVOPVFGFPTNVSLEWVNY-----IDIDELNGKLTTHCVLNRWR 106
      |  || | : :|| | | :|| | ||:| | |

```

```

0Y      44  KNAVKAIDRLHAGLFNTYSDVOQPVFGGPT--NVSLEMVVYI-DIDELNCKLTTHCW 100
      : | | : | : | | : | : | : | : | : | : | : | : | : | : | : | : |

```


GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: July 18, 2002, 08:50:47 ; Search time 79.64 Seconds
(without alignments)
615.062 Million cell updates/sec

Title: US-09-732-680A-2
Perfect score: 2311
Sequence: 1 MTTTPRIKAPVSGPGLPLL.....RISFVSFLAILAIRCSV 441

Scoring table:
BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 747574 seqs, 11073796 residues

Total number of hits satisfying chosen parameters: 747574

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :
A.Geneseq_032802:*
1: /SIDSI/gcgdata/geneseq/genesep-emb1/AA1980.DAT:*
2: /SIDSI/gcgdata/geneseq/genesep-emb1/AA1981.DAT:*
3: /SIDSI/gcgdata/geneseq/genesep-emb1/AA1982.DAT:*
4: /SIDSI/gcgdata/geneseq/genesep-emb1/AA1983.DAT:*
5: /SIDSI/gcgdata/geneseq/genesep-emb1/AA1984.DAT:*
6: /SIDSI/gcgdata/geneseq/genesep-emb1/AA1985.DAT:*
7: /SIDSI/gcgdata/geneseq/genesep-emb1/AA1986.DAT:*
8: /SIDSI/gcgdata/geneseq/genesep-emb1/AA1987.DAT:*
9: /SIDSI/gcgdata/geneseq/genesep-emb1/AA1988.DAT:*
10: /SIDSI/gcgdata/geneseq/genesep-emb1/AA1989.DAT:*
11: /SIDSI/gcgdata/geneseq/genesep-emb1/AA1990.DAT:*
12: /SIDSI/gcgdata/geneseq/genesep-emb1/AA1991.DAT:*
13: /SIDSI/gcgdata/geneseq/genesep-emb1/AA1992.DAT:*
14: /SIDSI/gcgdata/geneseq/genesep-emb1/AA1993.DAT:*
15: /SIDSI/gcgdata/geneseq/genesep-emb1/AA1994.DAT:*
16: /SIDSI/gcgdata/geneseq/genesep-emb1/AA1995.DAT:*
17: /SIDSI/gcgdata/geneseq/genesep-emb1/AA1996.DAT:*
18: /SIDSI/gcgdata/geneseq/genesep-emb1/AA1997.DAT:*
19: /SIDSI/gcgdata/geneseq/genesep-emb1/AA1998.DAT:*
20: /SIDSI/gcgdata/geneseq/genesep-emb1/AA1999.DAT:*
21: /SIDSI/gcgdata/geneseq/genesep-emb1/AA2000.DAT:*
22: /SIDSI/gcgdata/geneseq/genesep-emb1/AA2001.DAT:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2311	100.0	441	22	ABB68475
2	2311	100.0	441	22	AAB86218
3	420	18.2	627	18	AAW11825
4	418	18.1	627	18	AAW11824
5	416	18.0	529	15	AAW09023
6	416	18.0	529	15	AAW44155
7	416	18.0	529	16	AAW73966
8	416	18.0	529	18	AAW09021
9	412	17.8	450	22	AA67161
10	409	17.7	450	22	AAU00403
11	409	17.7	450	22	AAE00280

12	405.5	17.5	457	6	AAW50361
13	405.5	17.5	457	21	AAW56386
14	404.5	17.5	433	22	AAU00402
15	402.5	17.4	519	22	ABW62694
16	399.5	17.3	437	22	AAW6151
17	398.5	17.2	468	18	AAW09024
18	397.5	17.2	502	22	AAW51021
19	397	17.2	479	17	AAW89358
20	395.5	17.1	457	22	AAU00404
21	394	17.0	494	18	AAW09018
22	392.5	17.0	436	11	AAW06236
23	391	16.9	458	18	AAW09020
24	391	16.9	458	22	AAE12775
25	391	16.9	538	22	ABG27365
26	391	16.9	627	15	AAW44152
27	383.5	16.6	501	22	AAW78851
28	383.5	16.6	506	22	ABW11811
29	383.5	16.6	506	22	AAW79835
30	383	16.6	456	11	AAW06255
31	381	16.5	456	11	AAW06254
32	378.5	16.4	461	5	AAW40417
33	378	16.4	479	22	AAE12823
34	377.5	16.3	449	17	AAW86421
35	375.5	16.2	449	14	AAW35057
36	374	16.2	502	22	AAW50016
37	373	16.1	502	15	AAW44153
38	373	16.1	502	18	AAW09025
39	373	16.1	502	19	AAW69216
40	373	16.1	502	21	AAW24088
41	373	16.1	502	22	AAW82690
42	373	16.1	502	22	AAW50012
43	372	16.1	502	22	AAW50017
44	371	16.1	502	22	AAW50015
45	370.5	16.0	504	18	AAW09022

ALIGNMENTS

RESULT 1
ID ABB68475 standard; Protein: 441 AA.
XX
AC ABB68475;
XX
DE 26-MAR-2002 (first entry)
XX
DE Drosophila melanogaster polypeptide SEQ ID NO 32217.
XX
KW Drosophila: developmental biology; cell signalling; Insecticide;
KW pharmaceutical.
XX
OS Drosophila melanogaster.
XX
PN WO200171042-A2.
XX
PD 27-SEP-2001.
XX
PF 23-MAR-2001; 2001WO-US09231.
XX
PR 23-MAR-2000; 2000US-191637P.
XX
PR 11-JUL-2000; 2000US-0614150.
XX
PA (PEKE) PE CORP NY.
XX
PI Venter JC, Adams M, Li PWD, Myers EW;
XX
DR WPI; 2001-656860/75.
XX
DR N-PSDB; ABL12578.
XX
PT New isolated nucleic acid detection reagent for detecting 1000 or more
PT genes from Drosophila and for elucidating cell signalling and cell-cell
interactions -

XX Disclosure; SEQ ID NO 32217; 21pp + Sequence Listing; English.
PS
XX
CC The invention relates to an isolated nucleic acid detection reagent
CC capable of detecting 1000 or more genes from *Drosophila*. The invention is
CC useful in developmental biology and in elucidating cell signalling and
CC cell-cell interactions in higher eukaryotes for the development of
CC insecticides, therapeutics and pharmaceutical drugs. The invention
CC discloses genomic DNA sequences (AB101840-AB116175), expressed DNA
CC sequences (AB101840-AB116175) and the encoded proteins
CC (AB101840-AB116175).
CC The sequence data for this patent did not form part of the printed
CC specification, but was obtained in electronic format directly from WIPO
CC at ftp.wipo.int/pub/published_pcl_sequences.
XX
XX Sequence 441 AA:
SQ
Query Match 100.0%; Score 2311; DB 22; Length 441;
Best Local Similarity 100.0%; Pred. No. 5.1e-227;
Matches 441; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 MTTTPKIKAPVSGPGLLLQMLMGLTSPGATATADPKNAVKAIDRLHGLFT 60
Db 1 mtttprkikavsgpgrlllqmlmgllmgltsvpgatacadpknnavkaidrlhaglft 60
QY 61 NYDSVQPVFGPPTNVSLKLVVYIDIDELNGKLTTHCWLNRWDERVWQPSQYDNI 120
Db 61 nydsdvpgvfgprrptnslkvvlyididelnkltthcwnlrwdeervwqpsqydni 120
QY 121 TQITLKSEVWTPQITLFGNDEGGLMAETQYTLSDGHFRMPRPAYVTAYCELMLNMPH 180
Db 121 tqitlksewtpqitlfnngdegglmaetqytlshdghfrmprrpayvtaycelmlnmp 180
QY 181 DKOSCKIKISWGLKVVLPENGTARGESLDHDDLVQSPKMEIVDSRAHFVSQDYGYMEY 240
Db 181 dkgsckikiswglkvvlpengtargesldhddlvgspekwivdsrahfvsqdygymey 240
QY 241 TLTAQRSSMYTAVIYTPASCIYVLLASAFWLPDHMGKEKIMINGLLIYIAAFMYFAQ 300
Db 241 tltagrssmytavlytpasciyvllalsafwlpdhmgkekimgllliyiaafmyfaq 300
QY 301 LLPVLSNNTPLVYIFVSTSLIYLSVSTIVEVLYLATGKHRRLPALRKLHGLGTW 360
Db 301 llpvlsmntplvlyifvstsllylsvstivevlylatgkhrkrrlpalkllhghlgtw 360
QY 361 LLLSVFTTGESQAEKTKEMDEHPYEADQESSPLGINTEHVGAKANOFDMALLATAV 420
Db 361 lllsvfttgesqaeektkemdenhyeadeqessplginthevgakanofdmallata 420
QY 421 DRISFVSFLAFLILAIRCSV 441
Db 421 drisfvsflaflilaircsv 441
RESULT 2
ID AAB86218 standard; Protein: 441 AA.
XX
AC AAB86218:
XX
DT 31-AUG-2001 (first entry)
XX
DE D. melanogaster acetylcholine receptor beta subunit Db3 protein.
XX
KM Acetylcholine receptor beta subunit; Db3; fruitfly; insecticidal;
XX transgenic; modulator; plant-protection agent; medicine.
XX
OS *Drosophila melanogaster*.
XX
PN EP1106689-A2.
XX
PD 13-JUN-2001.

XX
XX 28-NOV-2000; 2000EP-0125300.
XX
XX 10-DEC-1999; 99DE-1059582.
XX
XX (FARB) BAYER AG.
XX
XX Adamczewski M, Methfessel C, Schulte T;
XX WPI: 2001-376906/40.
XX DR N-PSDB; AAH20992.
XX
XX New nucleic acid encoding *Drosophila* acetylcholine receptor subunit,
XX useful in screening for plant protection agents and pharmaceuticals
XX
XX Claim 9; Page 13-15; 18pp; German.
XX
XX This invention describes a novel nucleic acid (I) which encodes a
XX *Drosophila* melanogaster acetylcholine receptor beta subunit, Db3 which
XX has insecticidal activity. (I), related host cells, polypeptides, ACR,
XX antibodies, transgenic invertebrates and regulatory regions, are used to
XX identify agents that modulate ACR activity, potentially useful as
XX plant-protection agents (e.g. insecticides) or pharmaceuticals for human
XX or veterinary medicine. This sequence represents the *Drosophila* ACR
XX melanogaster acetylcholine receptor beta subunit described in the method
XX of the invention.
SQ
Sequence 441 AA:
SQ
Query Match 100.0%; Score 2311; DB 22; Length 441;
Best Local Similarity 100.0%; Pred. No. 5.1e-227;
Matches 441; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 MTTTPKIKAPVSGPGLLLQMLMGLTSPGATATADPKNAVKAIDRLHGLFT 60
Db 1 mtttprkikavsgpgrlllqmlmgllmgltsvpgatacadpknnavkaidrlhaglft 60
QY 61 NYDSVQPVFGPPTNVSLKLVVYIDIDELNGKLTTHCWLNRWDERVWQPSQYDNI 120
Db 61 nydsdvpgvfgprrptnslkvvlyididelnkltthcwnlrwdeervwqpsqydni 120
QY 121 TQITLKSEVWTPQITLFGNDEGGLMAETQYTLSDGHFRMPRPAYVTAYCELMLNMPH 180
Db 121 tqitlksewtpqitlfnngdegglmaetqytlshdghfrmprrpayvtaycelmlnmp 180
QY 181 DKOSCKIKISWGLKVVLPENGTARGESLDHDDLVQSPKMEIVDSRAHFVSQDYGYMEY 240
Db 181 dkgsckikiswglkvvlpengtargesldhddlvgspekwivdsrahfvsqdygymey 240
QY 241 TLTAQRSSMYTAVIYTPASCIYVLLASAFWLPDHMGKEKIMINGLLIYIAAFMYFAQ 300
Db 241 tltagrssmytavlytpasciyvllalsafwlpdhmgkekimgllliyiaafmyfaq 300
QY 301 LLPVLSNNTPLVYIFVSTSLIYLSVSTIVEVLYLATGKHRRLPALRKLHGLGTW 360
Db 301 llpvlsmntplvlyifvstsllylsvstivevlylatgkhrkrrlpalkllhghlgtw 360
QY 361 LLLSVFTTGESQAEKTKEMDEHPYEADQESSPLGINTEHVGAKANOFDMALLATAV 420
Db 361 lllsvfttgesqaeektkemdenhyeadeqessplginthevgakanofdmallata 420
QY 421 DRISFVSFLAFLILAIRCSV 441
Db 421 drisfvsflaflilaircsv 441
RESULT 3
ID AAW11825 standard; Protein: 627 AA.
XX
AC AAW11825:
XX
XX

```

DT 05-OCT-1997 (first entry)
XX Alpha4 subunit of mutant nAChR.
XX nAChR; mutation; autosomal dominant nocturnal frontal lobe epilepsy;
XX ADNFLE; neuronal acetylcholine receptor.
XX Homo sapiens.
XX
XX Key Location/Qualifiers
XX Misc-difference 280 /note= "Ser 248 Phe is the ADNFLE-
XX FT inducing mutation (see AAM11824)"
XX
XX A09656247-A.
XX
XX 09-JAN-1997.
XX
XX 28-JUN-1996; 96AU-0056247.
XX
XX 28-JUN-1995; 95AU-0003840.
XX
XX (UYBO-) UNIV BONN.
XX (UYME ) UNIV MELBOURNE.
XX (WOME-) WOMEN'S & CHILDREN'S HOSPITAL.
XX
XX Berkovic SF, Mulley JC, Phillips HA, Propping PJ;
XX Scheffer IE, Steinleinok, Sutherland GR, Wallace RH;
XX
XX WPI: 1997-100506/10.
XX
XX N-PSDB; AAT59528.
XX
XX CHRNA4 gene encodes neuronal nicotinic acetyl:choline receptor
XX fragment - used in diagnosis of autosomal dominant nocturnal frontal
XX lobe epilepsy
XX
XX Disclosure: Fig 13; 20pp; English.
XX
XX The DNA sequence of the normal nAChR is given in AAT59527,
XX the DNA sequence of the ADNFLE-inducing mutant is given in AAT59528.
XX Mutations at codon 248, which have the effect of replacing
XX serine by another amino acid (e.g. phenylalanine) in the sixth
XX amino acid of the transmembrane domain 2 (M2) of the alpha4
XX subunit of nAChR have been found to be associated with ADNFLE.
XX The primers given in AAT59529 and AAT59530 were used in the
XX amplification of part of exon 5.
XX
XX Sequence 627 AA:
SQ

```

Query Match 18.2%; Score 420; DB 18; Length 627;
 Best Local Similarity 29.8%; Pred. No. 1.2e-33;
 Matches 102; Conservative 67; Mismatches 139; Indels 34; Gaps 9;

```

OY 11 VSGPGILLQMLMGMLTSGVPGATATADPKNANVAKALRLHAGLTNTYDSVPYF 70
   :|||:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|
DB 3 1ggpgaprlpplllllgtgllraashvet-----rahaerlllklifsgynksrpy- 55
OY 71 OGPTPNVSLSENVVY-----IDIDELNGKLTTHCWLMLRRRDEBRWQPSQYDNTIOI 123
   :|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|
DB 56 ----anisdvnlvtrfglsiaqlidvdekngmttnvwvqgwhdyklrwdpadyenvtai 111
OY 124 TLKSSRWVPTQTLFNGDGGGLMAE--NQVTLSDHGFRFMRPAAVYTAACEINMLNMPHD 181
   :|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|
DB 112 rlpseilwrlpdlivynnaagdfavthltkahlfhdgrvqwrppatlyksscsidvtftrd 171
OY 182 KOSCKLIKIGS--GLKVVLPENGTARGESLDHDDLVSQPEWEIVDSRAHFVSODY----- 234
   :|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|
DB 172 gqncemkfgswtydkakdl-vnmhsr---vqgldtwegevwivdavgtyntiryecca 227
OY 235 --YGTMEVYTLAQRSSMTAVIYTPASCIIVILASAFULPRMHGGEKTMINGLLIIVIA 292
   :|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|
DB 228 e1ypdltyafavtrirp1rpflytlnlilpccilscitlvivfy1psec-gekiltcltfvllsit 286

```

```

OY 293 AFLMEFAQLPLVLSNNPLVWIFYSTSLYLSTIVLEVL 334
   ||:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|
DB 287 vlllitleipstslvpi1gey1lftm1fv1slv1cvfvl 328

```

RESULT 4
 AAM11824
 ID AAM11824 standard; Protein: 627 AA.
 XX
 XX AAM11824;
 AC
 XX
 DT 05-OCT-1997 (first entry)
 XX
 XX Alpha4 subunit of normal nAChR.
 XX
 XX nAChR; mutation; autosomal dominant nocturnal frontal lobe epilepsy;
 XX ADNFLE; neuronal acetylcholine receptor.
 XX
 XX Homo sapiens.
 XX
 XX Key Location/Qualifiers
 XX Misc-difference 280 /note= "Ser 248 Phe is the ADNFLE-
 XX FT inducing mutation (see AAM11825)"
 XX
 XX A09656247-A.
 XX
 XX 09-JAN-1997.
 XX
 XX 28-JUN-1996; 96AU-0056247.
 XX
 XX 28-JUN-1995; 95AU-0003840.
 XX
 XX (UYBO-) UNIV BONN.
 XX (UYME) UNIV MELBOURNE.
 XX (WOME-) WOMEN'S & CHILDREN'S HOSPITAL.
 XX
 XX Berkovic SF, Mulley JC, Phillips HA, Propping PJ;
 XX Scheffer IE, Steinleinok, Sutherland GR, Wallace RH;
 XX
 XX WPI: 1997-100506/10.
 XX
 XX N-PSDB; AAT59527.
 XX
 XX CHRNA4 gene encodes neuronal nicotinic acetyl:choline receptor
 XX fragment - used in diagnosis of autosomal dominant nocturnal frontal
 XX lobe epilepsy
 XX
 XX Disclosure: Fig 13; 20pp; English.
 XX
 XX The DNA sequence of the normal nAChR is given in AAT59527,
 XX the DNA sequence of the ADNFLE-inducing mutant is given in AAT59528.
 XX Mutations at codon 248, which have the effect of replacing
 XX serine by another amino acid (e.g. phenylalanine) in the sixth
 XX amino acid of the transmembrane domain 2 (M2) of the alpha4
 XX subunit of nAChR have been found to be associated with ADNFLE.
 XX The primers given in AAT59529 and AAT59530 were used in the
 XX amplification of part of exon 5.
 XX
 XX Sequence 627 AA:
 SQ

Query Match 18.1%; Score 418; DB 18; Length 627;
 Best Local Similarity 29.8%; Pred. No. 1.9e-33;
 Matches 102; Conservative 67; Mismatches 139; Indels 34; Gaps 9;

```

OY 11 VSGPGILLQMLMGMLTSGVPGATATADPKNANVAKALRLHAGLTNTYDSVPYF 70
   :|||:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|
DB 3 1ggpgaprlpplllllgtgllraashvet-----rahaerlllklifsgynksrpy- 55
OY 71 OGPTPNVSLSENVVY-----IDIDELNGKLTTHCWLMLRRRDEBRWQPSQYDNTIOI 123
   :|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|
DB 56 ----anisdvnlvtrfglsiaqlidvdekngmttnvwvqgwhdyklrwdpadyenvtai 111

```


XX PS Disclosure; Page 68-70; 99pp; English.
 XX CC The present sequence represents a human neuronal nicotinic acetylcholine
 CC receptor (nAChR) subunit. The cells expressing the alpha and/or beta
 CC nAChR subunits may be used in a method of screening compounds to
 CC identify any which modulate the activity of human neuronal nAChR.
 CC Subunit specific antibodies may be used to monitor the distribution
 CC and expression density of various subunits in normal vs diseased brain
 CC tissues. Testing of single receptor subunits or specific receptor
 CC subunit combinations with a variety of potential agonists or antagonists
 CC provides information with respect to the function and activity of the
 CC individual subunits and should lead to the identification and design of
 CC compounds that are capable of very specific interaction with one or
 CC more receptor subtypes. The resulting drugs should exhibit fewer
 CC unwanted side effects than drugs identified e.g. screening with cells
 CC that express a variety of subtypes.

XX S0 Sequence 529 AA;

Query Match 18.0%; Score 416; DB 15; Length 529;
 Best Local Similarity 27.6%; Pred. No. 2.3e-33;
 Matches 109; Conservative 75; Mismatches 153; Indels 58; Gaps 12;

OY 13 GPGELPILQLM-MGMLMGITSVPGATATADPKNA-----NYKALDR 53
 DB 2 GPSCPVILSFCKLSLWLLIPAGGEEKRPPTAPGDLSPSPALPGGSHETEDR 61
 OY 54 LHAGLFTNYDSVQVDFQCTPTNVSLEMYVY-----IDIDELNGKLTTHCWLNLNR 106
 DB 62 Lfkhlfirgynwarpv-----pntsdvivrfglsiaqlldvdeknqmttnvwlkqews 116
 OY 107 DEERYWQSQVDNITQITLKSEVWTPQITLFGNGDEGGLMAETQYTLSH---DGHFRMP 163
 DB 117 dyklirmpadfgnltslvpssemilwipdivlynadge-favltmktkahlstgtvhwvp 175
 OY 164 PAVYTAVCLELNLNMPHDKOSCKLRIGSM---GLKVLPENGTARGESLDHDLVQSPDW 220
 DB 176 palysscsidvtfpfdqgncmkfgswtydkakidleg---meqtdldkdyesgew 231
 OY 221 EIVDSRAHFVSODY-----YGYMEYTLTAQRSSMYTAIVYTPASCIYIALLSAFWLPP 274
 DB 222 alvatgynskkydcaaeiypdvgyafvtrrplpfylinllipclliscslvlfyips 291
 OY 275 HMGEEKIMINGLLIIVIAFLMYFAQLLPVLSNNTPLVVIFYSTSLVSTIVEVLV 334
 DB 292 dc-geklfclslvllstlvtlftelipstslvpligeyllftmifvltslvltvfl 350
 OY 335 YLATGKHKR-----RLPEALRKRLHGHIGTWTLLS 364
 DB 351 nv-----hnrpsptmhpwvrgallgcprvwlmm 381

RESULT 7

AAR73966 ID AAR73966 standard; Protein: 529 AA.

XX AC AAR73966;

XX DT 30-NOV-1995 (first entry)

XX DE Alpha 2 subunit of a human neuronal nicotinic acetylcholine receptor.

XX KW Human nAChR; neuronal nicotinic acetylcholine receptor;

XX RM neurotransmitter.

XX OS Homo sapiens.

XX PN W09513299-A.

XX PD 18-MAY-1995.

PF 08-NOV-1994; 94WO-US12859.

XX 08-NOV-1993; 93US-0149503.

XX PA (SALK) SALK INST BIOTECHNOLOGY IND ASSOC.

XX PI Elliott RJ, Ellis SB, Harpold MW;

XX WPI; 1995-194036/25.

XX N-PSDB; AA090387.

XX PT New human neuronal nicotinic acetylcholine receptor alpha2 subunit DNA
 PT used to develop prods. for detection, diagnosis and therapy and for
 PT modulating activity

XX PS Disclosure; Page 43-46; 54pp; English.

XX CC DNA encoding the human nAChR alpha2 subunit was isolated from a
 CC human thalamus tissue cDNA library using corresp. rat cDNA. The
 CC insert of one clone obtd. was ligated with the insert of another
 CC clone to generate a full-length alpha 2 subunit cDNA. The DNA can be
 CC used to identify function nAChRs. Cells contg. the DNA can be used
 CC for screening to identify cpds. which modulate the activity of human
 CC nAChRs. The human nAChR alpha 2 subunit can be used to product
 CC antibodies which can be used in immunohistochemistry, diagnosis and
 CC therapy. The nucleic acids can be used for analysing disease states
 CC and creating animal models.

XX S0 Sequence 529 AA;

Query Match 18.0%; Score 416; DB 16; Length 529;
 Best Local Similarity 27.6%; Pred. No. 2.3e-33;
 Matches 109; Conservative 75; Mismatches 153; Indels 58; Gaps 12;

OY 13 GPGELPILQLM-MGMLMGITSVPGATATADPKNA-----NYKALDR 53
 DB 2 GPSCPVILSFCKLSLWLLIPAGGEEKRPPTAPGDLSPSPALPGGSHETEDR 61
 OY 54 LHAGLFTNYDSVQVDFQCTPTNVSLEMYVY-----IDIDELNGKLTTHCWLNLNR 106
 DB 62 Lfkhlfirgynwarpv-----pntsdvivrfglsiaqlldvdeknqmttnvwlkqews 116
 OY 107 DEERYWQSQVDNITQITLKSEVWTPQITLFGNGDEGGLMAETQYTLSH---DGHFRMP 163
 DB 117 dyklirmpadfgnltslvpssemilwipdivlynadge-favltmktkahlstgtvhwvp 175
 OY 164 PAVYTAVCLELNLNMPHDKOSCKLRIGSM---GLKVLPENGTARGESLDHDLVQSPDW 220
 DB 176 palysscsidvtfpfdqgncmkfgswtydkakidleg---meqtdldkdyesgew 231
 OY 221 EIVDSRAHFVSODY-----YGYMEYTLTAQRSSMYTAIVYTPASCIYIALLSAFWLPP 274
 DB 222 alvatgynskkydcaaeiypdvgyafvtrrplpfylinllipclliscslvlfyips 291
 OY 275 HMGEEKIMINGLLIIVIAFLMYFAQLLPVLSNNTPLVVIFYSTSLVSTIVEVLV 334
 DB 292 dc-geklfclslvllstlvtlftelipstslvpligeyllftmifvltslvltvfl 350
 OY 335 YLATGKHKR-----RLPEALRKRLHGHIGTWTLLS 364
 DB 351 nv-----hnrpsptmhpwvrgallgcprvwlmm 381

RESULT 8

AAW09021 ID AAW09021 standard; Protein: 529 AA.

XX AC AAW09021;

XX DT 09-APR-1997 (first entry)

XX DE Neuronal nicotinic acetylcholine receptor alpha-2 subunit.

CC inflammatory disorders. They are also useful for treating and
 CC diagnosing disorders involving the spleen (e.g., Hodgkin disease,
 CC Niemann-Pick disease), lung (e.g., chronic bronchitis), colon
 CC (cirrhosis), uterus and endometrium (e.g., endometrial stromal tumours),
 CC brain (e.g., ischemia), T-cells (e.g., Sjogren syndrome), skin (lupus
 CC erythematosus), haematopoietic stem cells (e.g., Alzheimer's disease),
 CC heart (e.g., myocardial infarction), blood vessels (e.g., Kawasaki
 CC syndrome), red cells (e.g., anemias), disorders involving thymus,
 CC B-cells, kidney (e.g., glomerulonephritis), disorders involving breast,
 CC testis, epididymis, prostate, thyroid (e.g., Graves disease), disorders
 CC involving skeletal muscle (e.g., tumour), pancreas (e.g., pancreatitis),
 CC small intestine (e.g., celiac sprue), disorders related to reduced
 CC platelet number and ovary.

SO Sequence 450 AA:

Query Match 17.8%; Score 412; DB 22; Length 450;
 Best Local Similarity 27.9%; Pred. No. 4.6e-33;
 Matches 127; Conservative 78; Mismatches 188; Indels 62; Gaps 18;

QY 24 MGLMLGTSVPGATATADPKANVKALDRHAGFTNDSVQVYFQCTPT-NVSLKAV 82
 DB 10 IGIILILF--IpaecIgaegrial-----klfrdIfanytsalrpadtdqglntvleat 62
 QY 83 VT-YIDIDELNGKLTTHCMLNLRMRDEERWQPSQYDNTOTLKSEVWFOITLFGND 141
 DB 63 IsgitdmderngvltlylwirgewtdaylrwbpnagygidaIrpslswrpdvlylnka 122
 QY 142 EGGI--MAETQVTLSDHGFRMRPRAVYATYACELMMLMWRHDQSCKLKIGSM--GLKV 196
 DB 123 daqpgsasltvnlrhdgavtrwdapaIttrscrvdaafpfdaqngcltfgswthgngql 182
 QY 197 -VLPENGTARGSLDHDLVQSPWEIYDSRA-----HFVSQDYGYGMEYTLTKQRSS 249
 DB 183 dvTP-----rgaaaslaafvenvewrVlqmparrvlygcscsepydvtflllrrraa 237
 QY 250 MYTAIVTPASCIYVILASAFWLPRHMGCEKIMINGLITVIAAFYAOQLPVLNNNT 309
 DB 238 aycvclllpvcvllslaplafhlpad-sgekvisgtvtlaItvfqlllaesmp-aeav 295
 QY 310 PLVVIYSTSLYSVSTIVEVLVYL-ATGKHKRRPEALRKLLHGLGTWLLSVFST 368
 DB 296 pllgkyymatmvtfstaltlllmhlycgsrvrpywparalllghnarglcv---re 352
 QY 369 TGESQAE-KTKEMDHPYFEADESSPLG-----INHEVPGAKKNQF-- 411
 DB 353 rgepcgqsrpelpsp-qspegagppapgcpcpcrcIcrgaalhhv---atlantfrs 408
 QY 412 -----DNALLATAVDRISEVFSLSAFLIAL 437
 DB 409 hraagrchedwkrIarvmrfflalfismaIvmsl 443

RESULT 10
 AAU00403
 ID AAU00403 standard; Protein: 450 AA.

AC AAU00403;

DT 12-SEP-2001 (first entry)

DE Human alpha nicotinic acetylcholine receptor amino acid sequence.

KW Human; alpha nicotinic acetylcholine receptor; alpha10AChR; asthma;
 KW chronic obstructive lung disease; adult respiratory distress syndrome;
 KW sepsis; rheumatoid arthritis; osteo-arthritis; Crohn's disease;
 KW inflammatory bowel disorder; psoriasis; myasthenia gravis; schizophrenia;
 KW epilepsy; Parkinson's disease; Alzheimer's disease; Tourette's syndrome;
 KW chronic pain; nicotine addiction; hormonal deficiency.

OS Homo sapiens.

XX
 XX

PN WO200119973-A2.

XX 22-MAR-2001.

XX 14-SEP-2000; 2000WO-EP09115.

XX 15-SEP-1999; 99US-0153948.

PR 02-FEB-2000; 2000GB-0002431.

XX (JANC) JANSSEN PHARM NV.

XX Yon JR, Grantham CJ, Groot-Kormelink PJ;

DR WPI: 2001-244795/25.

DR N-PSDB: AAS01280.

PT Novel human alpha nicotinic acetylcholine receptor (alpha10AChR) useful
 PT for identifying modulator compounds for treating psoriasis,
 PT schizophrenia, asthma, Alzheimer's disease, chronic pain, and nicotine
 PT addition

XX Claim 4: Fig 6; 72pp; English.

CC The sequence represents the amino acid sequence of human alpha
 CC nicotinic acetylcholine receptor (Alpha10AChR). The nucleic acid and
 CC proteins can be used in identifying modulator compounds useful in the
 CC treatment of asthma, chronic obstructive lung disease, acute adult
 CC respiratory distress syndrome, sepsis, rheumatoid and osteo-arthritis,
 CC inflammatory bowel disorder, Crohn's disease and psoriasis, myasthenia
 CC gravis, schizophrenia, epilepsy, Parkinson's disease, Alzheimer's
 CC disease, Tourette's syndrome, chronic pain and nicotine addiction.
 CC The nucleic acid can be used in a test for detecting alpha10AChR encoding
 CC sequences in the human body or tissues or samples obtained from humans.
 CC They are also useful in tissue distribution studies in human health and
 CC disease. Overexpression of the protein in cell lines may be used to
 CC produce membrane preparations bearing the protein for ligand binding
 CC studies. The protein is useful as an immunogen to obtain specific
 CC antibodies which are useful in purification and manipulation of
 CC polypeptides. The modulator compounds may be used to modulate the
 CC ion channel activity of alpha10AChR. The modulator compounds also have
 CC therapeutic value in diseases caused by hormonal deficiency or
 CC over-production of hormones.

XX Sequence 450 AA:

SO

Query Match 17.7%; Score 409; DB 22; Length 450;
 Best Local Similarity 27.9%; Pred. No. 9.4e-33;
 Matches 127; Conservative 78; Mismatches 188; Indels 62; Gaps 18;

QY 24 MGLMLGTSVPGATATADPKANVKALDRHAGFTNDSVQVYFQCTPT-NVSLKAV 82
 DB 10 IGIILILF--IpaecIgaegrial-----klfrdIfanytsalrpadtdqglntvleat 62
 QY 83 VT-YIDIDELNGKLTTHCMLNLRMRDEERWQPSQYDNTOTLKSEVWFOITLFGND 141
 DB 63 IsgitdmderngvltlylwirgewtdaylrwbpnagygidaIrpslswrpdvlylnka 122
 QY 142 EGGI--MAETQVTLSDHGFRMRPRAVYATYACELMMLMWRHDQSCKLKIGSM--GLKV 196
 DB 123 daqpgsasltvnlrhdgavtrwdapaIttrscrvdaafpfdaqngcltfgswthgngql 182
 QY 197 -VLPENGTARGSLDHDLVQSPWEIYDSRA-----HFVSQDYGYGMEYTLTKQRSS 249
 DB 183 dvTP-----rgaaaslaafvenvewrVlqmparrvlygcscsepydvtflllrrraa 237
 QY 250 MYTAIVTPASCIYVILASAFWLPRHMGCEKIMINGLITVIAAFYAOQLPVLNNNT 309
 DB 238 aycvclllpvcvllslaplafhlpad-sgekvisgtvtlaItvfqlllaesmp-aeav 295
 QY 310 PLVVIYSTSLYSVSTIVEVLVYL-ATGKHKRRPEALRKLLHGLGTWLLSVFST 368
 DB 296 pllgkyymatmvtfstaltlllmhlycgsrvrpywparalllghnarglcv---re 352

Db	409	hnaaqrcdckkrlarwymdrfflafflffsmalvmsl	443
RESULT	12		
ID	AAP50361		
XX	AAP50361 standard; Protein: 457 AA.		
XX			
AC	AAP50361;		
XX			
DT	08-JAN-1992 (first entry)		
XX			
DE	Human acetyl choline receptor alpha.		
XX			
KW	hACR-alpha; myasthenia gravis; MG; neuromyopathy.		
XX			
OS	Homo sapiens.		
XX			
FH	Key	Location/Qualifiers	
FT	Active-site	42..47	
FT		/label= Epitope	
FT		/note= "Claim 1 (I)"	
FT		86..91	
FT	Active-site	/label= Epitope	
FT		/note= "Claim 1 (II)"	
FT		158..163	
FT	Active-site	/label= Epitope	
FT		/note= "Claim 1 (III)"	
FT		181..186	
FT	Active-site	/label= Epitope	
FT		/note= "Claim 1 (IV)"	
FT		198..203	
FT	Active-site	/label= Epitope	
FT		/note= "Claim 1 (V)"	
FT		354..359	
FT	Active-site	/label= Epitope	
FT		/note= "Claim 1 (VI)"	
FT		407..412	
FT	Active-site	/label= Epitope	
FT		/note= "Claim 1 (VII)"	
XX			
PN	JP60078996-A.		
XX			
PD	04-MAY-1985.		
XX			
PF	05-OCT-1983; 83JP-0186402.		
XX			
PR	05-OCT-1983; 83JP-0186402.		
XX			
PA	(MITU) MITSUBISHI CHEM IND KK.		
XX			
DR	WI; 1985-144120/24.		
XX			
DR	N-PSDB; AAN50415.		
XX			
PT	New peptide(s) useful in treatment of myasthenia gravis - also in		
XX	diagnosis, having acetyl:choline receptor-alpha like activities.		
PS	Disclosure; Flg 1; 13pp; Japanese.		
XX			
CC	Peptides derived from the hACR-alpha gene product have similar		
CC	immunogenicity to the gene product, and may be useful in the		
CC	treatment and diagnosis of myasthenia gravis caused by neuropathy in		
CC	autoimmune reaction to ACR.		
XX			
Sequence	457 AA:		

Query Match	17.5%	Score 405.5;	DB 6;	Length 457;
Best Local Similarity	26.9%;	Pred. NO. 2.2e-32;		
Matches 116;	Conservative 84;	Mismatches 165;	Indels 67;	Gaps 16;
53 RHAGLFFNYDDVQPV---FQGTPTNSLENAVYYIIDDELNGKLTTCWMLNKRDEE	109			

Dd		26	rIvakiLfkyssvvrvpedhgrqvventvgIql-IqIInvdenrqivtInvlIkqgwvdn	84
Oy		110	RwvQPSQDNtNQITLkSSEWTPPQITTFNGDEG--GLMAETQVTLSDHGFRMPAPAY	167
Dd	I	85	LkpnpddYggyvkKihIpsekiwrpdIvlnnadgdfaIvkfkcvllIgcghltwcpalf	144
Oy		168	TNYCELMNMNMFHDOSCKLTIGSW----GLKVILPENGTARESIDHDDLVDSPEMEY	223
Dd	:	145	kSyceIivhIfpfdeqcsmKIgtwcYdgsvvaInpes-----dqpdIsInfmesgeviX	199
Oy		224	DSRAHFVSODY-----GYMEYTELTAAQRSSMYRAYITYTSPASCIIVIALSAFMPPHM	276
Dd	S	200	eerygmKhstYecpdrtpIdltYhvmfrIprYifvwnlIipollsfItglvfYlpdt-	258
Oy		277	GGEKTMINGLLIIYAATFYAQLLPVLNSNPPLVIFYSTSLAVSVTIVEVLVL	336
Dd	I	259	sgekmItIsstvlslstlvflvIlveIlIPstcsavplIgikYmIlfemvfvlaSlItvIvI--	316
Oy		337	AGKKRR-----LRPALKLHGHCWTULLSVFTSGESQAERKKEDEHHY-BEAD	390
Dd	-	317	nuthrpstchmpmwrvkvf--IdtlpImfistmkrrprek--qdKklftedIdl	368
Oy		391	OESS-----PLG-----INHTEVPGA-----KANQ-----FDMLLATAYDR	422
Dd	:	369	sdlsgkpppppgfnfpIspIkhpevksaiegIyiaetmksdgeSmaaaekwyamymdh	428
Oy		423	ISFVSFSLAFLT	434
Dd	I	429	Illygfmlvcll	440
RESULT 13				
AAID	AY56386			
ID	AA56386	standard; Protein;	457 AA.	
XX	AA56386;			
DT	14-FEB-2000	(first entry)		
DE	Human acetylcholine receptor protein sequence.			
KM	Human; acetylcholine receptor; AChR; immune response; allergy;			
KW	antibody mediated disease; gene replacement therapy; T cell epitope;			
KW	dermatological; immunosuppressive; antiinflammatory; haemostatic;			
KM	antihaemic; antiallergic; antiasthmatic; antithyroid; antidiabetic;			
KW	autoimmune disease; allergic disease.			
XX	Homo sapiens.			
OS	Wo9930736-A2.			
PN	24-JUN-1999.			
PD	16-DEC-1998;	98WO-US26787.		
XX	16-DEC-1997;	97US-0991143.		
PR	(MINU) UNITV MINNESOTA.			
PA	Conti-Fine BM;			
P1	WPI; 2000-038343/03.			
XX	N-PSDB; AA238821.			
DR	Use of T cell epitope peptides for, e.g. preventing allergies -			
XX	Example 2; Page 219-221; 221pp; English.			
PS	The present invention describes a method for preventing or inhibiting an			
CC	indication of disease associated with aberrant, pathogenic or undesirable			
CC	antibody production, particularly autoimmune or allergic diseases. The			
CC	method comprises administering a T cell epitope peptide. The sequence of			
CC	the epitope peptide comprises an immunodominant T cell epitope sequence			

CC which is less than the sequence of the antigen. The antigen comprises
CC the immunodominant T cell epitope sequence. The method can be used to
CC specifically tolerate or down regulate the priming or activity of
CC antigen-specific T cells of a mammal. It can be used to prevent or
CC inhibit an indication or disease associated with antibody production to
CC an antigen such as an endogenous antigen, e.g. acetylcholine receptor,
CC insulin, growth hormone, factor VIII or factor IX, or an exogenous
CC antigen e.g. a fungal antigen, a plant antigen, an antigen of a domestic
CC cat or an antigen of a mite. It can be used to treat autoimmune diseases,
CC e.g. myasthenia gravis, systemic lupus erythematosus (SLE), Grave's
CC disease, autoimmune haemolytic anaemia, autoimmune thrombocytopaenia,
CC autoimmune asthma, cryoglobulinaemia, thrombotic thrombocytopenic purpura,
CC primary biliary sclerosis, pernicious anaemia or pemphigus. It can be
CC used for treating allergic diseases, e.g. allergic rhinitis, allergic
CC asthma, atopic dermatitis, allergic gastroenteropathy, anaphylaxis,
CC urticaria or angioedema. It can also be used in gene therapy for treating
CC a disease such as haemophilia or diabetes or an indication such as
CC adenosine deaminase deficiency, growth hormone deficiency, insulin
CC deficiency, factor IX deficiency or factor VIII deficiency. The present
CC sequence represents human acetylcholine receptor which is used in the
CC exemplification of the present invention.

SQ Sequence 457 AA;

Query Match	17.5%;	Score 405.5;	DB 21;	Length 457;
Best Local Similarity	26.9%;	Pred. No. 2.2e-32;		
Matches 116;	Conservative 84;	Mismatches 165;	Indels 67;	Gaps 16

```

0Y 53 RLHAGLFNYSODPV---FQGRPTNVLSEMYVTYIDIDELNOKLTHTCMLNRMDEE 109
Db 26 lrvaklfdydsavvrpdedhqvvevtgldl-qlnveevnqvltnvrlkqwdvyn 84
0Y 110 RVMOPSOYDNTTOITLLKSSEWTPQITLFGNDEC--GMAETQYTLSDHGHFRMPRAVY 167
Db 85 lkmpdddyguykklhipekltwpolvlynnagdfavtkvllgytvglltwptraif 144
0Y 168 TAYCELNLNMPHROKOSKLTGSM---GLKVLVPENGTAAGSLDHDLVQSPHEVY 223
Db 145 ksyceilythrpfdqegmsklgwcvgsvvalnpes----dqpolshmesgevvik 199
0Y 224 DSRANFVSQDY-----YGYMETYLLAQRSSMYTAVIYTPASCIYLLASAPRPHM 276
Db 200 eargwkshvetscoppdpyldityhfmqrjprlyfynvllpollfsflgvlgytbrtd- 258
0Y 277 GGEKTMINGLLIIVIAFLMTFAOLLPYLSNNPLVVIESTSLIYTSVSTIEVLYVL 336
Db 259 sgekmctlsayvlslitvlylivelipstsaavylldkymfltwvfaasllitvlyl- 316
0Y 337 ATGKHRR-----PPEARLKLHLHGLGWLLLSVFSTGESAQAKTKEMDENHP-EEADE 392
Db 317 --nchrspschmpnvrvkfv-----ldipnlmftcmkprprek---gdkkiftedid 366
0Y 391 QESS-----BLG-----INHTEPGA-----KANQ-----FDWALLTAADR 422
Db 369 sdlsqgkppppmgfhsplrlkhpvksalegklylaetcmksdgesnaaaekwyamwmdh 428
0Y 423 ISFVSFLAFLI 434
Db 429 lljgvfmlgvcll 440

```

RESULT	14
AAU00402	
ID	AAU00402 standard; Protein; 433 AA

AC	AAU00402;
XX	
DT	12-SEP-2001 (first entry)
XX	
DE	Mature human alpha nicotinic acetylcholine receptor amino acid sequence
XX	
Human:	alpha nicotinic acetylcholine receptor; alpha10AChr; asthma;

chronic obstructive lung disease; adult respiratory distress syndrome;
sepsis; rheumatoid arthritis; osteo-arthritis; Crohn's disease;
inflammatory bowel disorder; psoriasis; myasthenia gravis; schizophrenia;
epilepsy; Parkinson's disease; Alzheimer's disease; Tourette's syndrome;
chronic pain; nicotine addiction; hormonal deficiency.

Homo sapiens.
WO200119973-A2.

22-MAR-2001

14-SEP-2000; 2000WO-EP09115.

15-SEP-1999; 99US-0153948.

02-FEB-2000; 2000GB-0002431.

(JANC) JANSSEN PHARM NV.

Yon JR, Grantham CJ, Groot-Kormelink PJ,

WPI; 2001-244795/25.

N-PSDB; AAS01279.

Novel human alpha nicotinic acetylcholine receptor (alpha10AChR) useful for identifying modulator compounds for treating psoriasis, schizophrenia, asthma, Alzheimer's disease, chronic pain, and nicotine addiction -

Claim 4; Page 71-72; 72pp; English.

The sequence represents the amino acid sequence of mature human alpha nicotinic acetylcholine receptor (Alpha10AChR). The nucleic acid and proteins can be used in identifying modulator compounds useful in the treatment of asthma, chronic obstructive lung disease, acute adult respiratory distress syndrome, sepsis, rheumatoid and osteo-arthritis, inflammatory bowel disorder, Crohn's disease and psoriasis, myasthenia gravis, schizophrenia, epilepsy, Parkinson's disease, Alzheimer's disease, Tourette's syndrome, chronic pain and nicotine addiction. The nucleic acid can be used in a test for detecting alpha10AChR encoding sequences in the human body or tissues or samples obtained from humans. They are also useful in tissue distribution studies in human health and disease. Overexpression of the protein in cell lines may be used to produce membrane preparations bearing the protein for ligand binding studies. The protein is useful as an immunogen to obtain specific antibodies which are useful in purification and manipulation of polypeptides. The modulator compounds may be used to modulate the ion channel activity of alpha10AChR. The modulator compounds also have therapeutic value in diseases caused by hormonal deficiency or over-production of hormones.

Sequence 433 AA

Query Match	17.5%;	Score 404.5;	DB 22;	Length 433;
Best Local Similarity	28.2%;	Pred. No. 2.6e-32;		
Matches 120;	Conservative 72;	Mismatches 179;	Indels 55;	Gaps 16;

```
Oy      53 RlhAGLEfTNYDSdVQPYEgGfPT-NvSIEMwVT-YIDIdELNGKLTTHCWNLrWRkDEER 110
       :| | | | :|| | | | :| | | | :| | | | :| | | |
Db     15 klfdrI fanytsalrpvadtdgtlnvtlevtlsgidmdernqvltlylwirgetdayl 74
```

```

QY      111  WDPQSYDNITQITLESSEWWTPOITLNGDEGL--MAYQVILSHDGHRRMMPAVYT 168
      | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db      75  rwdpnaygqldairlpsslwpvpidvlynkadapggssstnvlrhdgavrwdaiptr 134

```

QY 169 AYCELNMLNWPBKQCKLIGSW--GLKV-VLPENGARGESLDHDDLVLQSPWEIVD 224

Db 135 sscrvdvaafpfdaghcgltfqswthgqhldvr-----rgaasladfvenvewrvlg 189

Oy 225 SRA-----HEVSQDYGYMEYTLTAORSSMYTAVITPASCIVIALALSAFWLPPHMG 278
| : | : ||| :|| : | | :|| ||| |
Db 190 mparrrvltgycgcsepydvttlllrrraaycnlllpvcvlslaplafnhpad-sg 248

us-09-732-680a-2.ray

OY 279 EXIMINGILLITVIAFAWMEFAOPLPLSNNTPLVFAVSTSLLYASTVTEVLVLYL-A 337
 Db 249 ekvsIgsvevllatltvfglllaesmp-aeavpllgkyymatmuvfstctlllmhly 307
 OY 338 TGRKRRRLPEALRKLLHGLGTWLLSVSTGESQAE-KTKEMDHPYEADQESSPL 396
 Db 308 cpgsvrppawaralllghlrglcw---rergpcgsgrrpelpsp-qspeggqppa 363
 OY 397 G-----INHEVPGAKRNOF-----DVALLATAVDRISFVSLSLA 431
 Db 364 gpcheprclcrgaallhvh--atlantlrshraqichedwKrlarvmdrflaifsm 420
 OY 432 FLILAI 437
 Db 421 alvmsl 426
 RESULT 15
 ABB62694
 ID ABB62694 standard; Protein; 519 AA.
 AC ABB62694;
 XX 26-MAR-2002 (first entry)
 DT
 XX Drosophila melanogaster polypeptide SEQ ID NO 14874.
 DE
 XX Drosophila: developmental biology; cell signalling; insecticide;
 KW pharmaceutical.
 XX
 OS Drosophila melanogaster.
 PN WO200171042-A2.
 XX
 PD 27-SEP-2001.
 XX
 PF 23-MAR-2001; 2001WO-US092331.
 XX
 PR 23-MAR-2000; 2000US-191637P.
 PR 11-JUL-2000; 2000US-0614150.
 XX
 PA (PEKE) PE CORP NY.
 PA
 PI Venter JC, Adams M, Li PWD, Myers EW;
 XX
 DR MPI: 2001-656860/75.
 DR N-PSDB; ABL06797.
 XX
 PT New isolated nucleic acid detection reagent for detecting 1000 or more
 PT genes from Drosophila and for elucidating cell signalling and cell-cell
 PT interactions -
 XX
 PS Disclosure: SEQ ID NO 14874; 21np + Sequence Listing; English.
 CC The invention relates to an isolated nucleic acid detection reagent
 CC capable of detecting 1000 or more genes from Drosophila. The invention is
 CC useful in developmental biology and in elucidating cell signalling and
 CC cell-cell interactions in higher eukaryotes for the development of
 CC insecticides, therapeutics and pharmaceutical drugs. The invention
 CC discloses genomic DNA sequences (ABLI0176-ABU30511), expressed DNA
 CC sequences (ABLI01840-ABLI6175) and the encoded proteins
 CC (ABBS7737-ABR72072).
 CC The sequence data for this patent did not form part of the printed
 CC specification, but was obtained in electronic format directly from WIPO
 CC at ftp.wipo.int/pub/published_pct_sequences.
 XX
 SQ Sequence 519 AA:

[illegible]

Search completed: July 18, 2002, 10:16:50
Job time: 5163 sec

Query Match	17.4%;	Score 402.5;	DB 22;	Length 519;
Best Local Similarity	25.5%;	Pred. No. 5.4e-32;		
Matches 123;	Conservative 92;	Mismatches 192;	Indels 75;	Gaps 16

Fri Jul 19 08:26:19 2002

us-09-732-680a-2.rag

